Social media, financial reporting opacity and return co-movement: Evidence from Seeking Alpha

Abstract

In this study we develop a model to analyse the interplay between social media coverage, financial reporting opacity and stock return co-movement. Our model predicts a negative association between social media coverage and co-movement, because social media coverage lowers the information acquisition and processing cost for investors and therefore facilitates the incorporation of firm-specific information into stock price, which leads to reduced co-movement. It is also predicted that such effect is more pronounced among firms with higher financial reporting opacity. Using data collected from Seeking Alpha, the largest social media platform providing "third-party generated" financial commentary and analysis in the US, we find results consistent with the predictions of the model. Our study has significant policy implications, because social media has become an increasingly important channel of information production and dissemination.

Keywords: social media; co-movement; Seeking Alpha; financial reporting opacity

JEL: G11, G12, G14

1. Introduction

In this study, we develop a model to analyze the interplay between social media coverage, financial reporting opacity and the extent to which stock return co-moves with industry and market return (co-movement). Higher information acquisition and processing costs may significantly reduce the number of informed investors, which in turn deters firm-specific information from being fully capitalised into stock prices (Grossman and Stiglitz, 1980), resulting in high co-movement between stock return and market/industry return. However, financial analysis written and posted by registered users on social media might have "global access" on the Internet and significantly decrease the information collection and processing costs, which facilitates the incorporation of more information, particularly firm-specific information, into stock price. We suggest that the coverage of public firms on social media plays at least two roles:

- Firstly, the coverage may help investors become familiar with the firm covered at a low cost, resulting in reduced information asymmetry and better investor awareness. Given that there are an increasing number of informed investors who can exploit their private information when they trade with uninformed counterparts, firm-specific information is likely to be incorporated into stock price to a greater extent;
- Secondly, prior research shows that contagion is an important type of social influence that takes the form of direct interaction with others in a network, and such influence can explain a wide range of economic activities (Bikhchandani *et al.*, 1998; Ivković and Weisbenner, 2007). In the financial market, interaction with other market participants is critical in the contagion of value relevant information, as people pay more attention to ideas or facts that are reinforced by interaction, ritual and symbols (Shiller, 1999).[1] In the Internet era, social media is becoming an important platform where individuals can

[1] For example, Shiller and Pound (1989) find that almost all investors who recently purchased a stock had their attention drawn to it through direct communication with others.

exchange ideas on investment with others, because it allows registered user not only to write and read articles, but also to post commentaries in response to an existing piece of information. Those who post commentaries may provide distinctive perspectives or alternative insights, and can suggest corrections or even point out flaws in the original article. Not unreasonably, an article followed by many commentaries is expected to attract more attention from a broad audience. Therefore, the coverage on social media enables more investors to better interact with others and understand the implication of released information, which results in such firm-specific information being fully impounded into stock price, leading to reduced return co-movement.

It is recognized that financial reports are an important source of firm-specific information that is widely used by investors (Bushman and Smith, 2001; Lambert, 2001). According to Jin and Myers (2006), (financial reporting) opacity, which represents the lack of information that precludes investors from determining the fair value of a firm, makes it easier for managers to conceal self-serving behavior such as rent seeking or asset diversion.[2] This implies that the opacity limits the flow of firm-specific information to the market, which leads to higher return co-movement. Under such circumstances, alternative information sources such as social media enable investors to access "third-party generated information" related to a firm. This suggests that social media has a stronger effect in facilitating the flow of firm-specific information to the market for firms with high financial reporting opacity, resulting in stock return of such firms being less co-moving with the market and industry returns. In section 2 we develop a model to formally derive propositions regarding the relation between social media coverage, financial reporting opacity and return co-movement.

[2] In an analytical study Bleck and Liu (2007) posit that opacity in financial reporting constrains shareholders' ability to distinguish bad projects from good projects at an early stage, leading to the continuation of bad projects over time and a higher risk of crash in asset prices. Such a prediction is supported by Hutton *et al.*, (2009), which report a positive association between crash risk and the extent firms engage in accrual-based earnings management.

To empirically test the predictions of our model, we design a computer program to automate the process of extracting all articles published on Seeking Alpha website between 2004 and 2014.[3] We focus on single-ticker articles that only provide information about one specific firm, and remove from our analysis all multiple-ticker articles that discuss more than one firm in one article. Then we focus on firms that have been covered by Seeking Alpha at least once in our sample period, and construct the coverage measure as the log of one plus the number of single-ticker Seeking Alpha articles for a firm during a specific year.

Our first measure of co-movement is stock price synchronicity, defined as the extent to which variation in firm-level stock return can be explained by market and industry returns (Durnev *et al.*, 2003; Piotroski and Roulstone, 2004; Crawford *et al.*, 2012). Following Roll (1988), we measure stock price synchronicity with adjusted R^2 from the market model regression to capture the extent to which stock price movement can be explained by both market and industry-wide information.[4] After a log-transformation, a lower synchronicity measure implies that market and industry returns can explain a smaller proportion of individual stock returns, suggesting that more firm-specific information has been capitalized into stock price. Consistent with our prediction, our results confirm that Seeking Alpha coverage is associated with lower synchronicity (lower level of co-movement). Such a result is economically significant, because for firms covered by Seeking Alpha, a one-standard deviation increase in Seeking Alpha coverage is associated with a 4.3% reduction in the synchronicity measure.[5]

Our second measure of co-movement is the percentage point of time-series Pearson

[3] Seeking Alpha, which was founded in 2004 by David Jackson, is a social media platform that publishes financial commentary and analysis. By the end of 2015 it had 4 million registered users, while more than 10,000 registered users contribute financial commentary and analysis. Submitted articles are reviewed by a panel and are subject to editorial changes, so the quality of published articles is expected to be high. There are 7 million average monthly unique visitors and 85 million average monthly page views. Seeking Alpha has a broad coverage of stocks, including 4,000 small and mid-cap firms.

[4] This measure of return co-movement (and its variation) has been widely used in previous research (i.e., Morck *et al*., 2000; Fernandes and Ferreira, 2008; Gul *et al*., 2011).

[5] The reduction is calculated as percentage change in co-movement before log transformation, which is the adjusted R^2 from regressing firm level stock return on market and industry returns.

correlation coefficient between weekly firm return and weekly market return (CORRE). Based on this measure we find consistent results that Seeking Alpha coverage is associated with lower co-movement. We further corroborate that our findings are robust to estimation using a matched sample based on a propensity score (PSM), a two-stage least square (2SLS) approach, and decomposing the synchronicity measure into a market return component and an industry return component. Results show that Seeking Alpha coverage is negatively associated with both the market return component and the industry return component. Taken together, our findings suggest that social media plays an important role in facilitating the flow of firm-specific information to the market, thus decreasing return co-movement.

Next, we investigate whether the association between social media coverage and return co-movement is more pronounced in firms with higher financial reporting opacity, because higher financial reporting opacity is expected to make it difficult for investors to obtain information from financial reports, thus making them increasingly reliant on alternative information source such as social media to access firm-specific information. Consistent with Hutton *et al.*, (2009), we use the average discretionary accruals calculated from the Francis *et al.*, (2005) model over a three-year period as the first proxy of financial reporting opacity, and find evidence supporting the prediction of our model that the effort of social media coverage on co-movement is more pronounced in firms with higher financial reporting opacity. Our inferences are qualitatively unchanged when we use analyst forecast dispersion as the second measure of financial reporting opacity, with the rationale that forecast dispersion might be larger for firms with higher opacity.

The contribution of our study is threefold. Firstly, we extend the literature on the capital market consequence of social media (Blankespoor *et al.*, 2014; Lee *et al.*, 2015) by developing a model to explicate the relation between social media coverage, financial reporting opacity and co-movement.[6] The predictions of the model are supported by data collected from Seeking Alpha. As a typical example of social media specializing in financial analysis, Seeking Alpha article, which is categorised as “third-party generated information”, substantially differs from a firm-initiated press release on Facebook or Twitter. Seeking Alpha articles are composed by registered users and independent parties (i.e. the editorial team of Seeking Alpha) will verify the quality of submission and the credentials of the author (i.e. name, address and contact information) before an article is eventually published.[7] Furthermore, unlike tweets, which are restricted to 140 characters until September 2017, Seeking Alpha articles can accommodate in-depth analysis of a firm, thus conveying valuable information with figures and numbers to validate the information. In particular, since January 2011 Seeking Alpha started paying each contributing author $10 per 1,000 page views on her article.[8] It is likely, therefore, that the information revealed in a published article on Seeking Alpha is of high quality, because the author has a financial incentive to publish articles with high credibility. If the information provided in an article proves to be misleading, the reputation of the author will be negatively affected, and any future article from the same author is unlikely to be published by Seeking Alpha. Similar incentives are absent from both Twitter and Facebook, which implies that Seeking Alpha constitutes a power setting to test our model.[9]

---

[6] Blankespoor *et al.*, (2014) find that firms can reduce information asymmetry by disseminating corporate news through Twitter, and Lee *et al.*, (2015) show that firms use social media such as Twitter to interact with investors to attenuate the negative stock market reaction to product recalls.

[7] According to an article written by Seeking Alpha CEO Ali Hoffmann and posted on its website on 10th April 2014, the editorial team of Seeking Alpha evaluates each submission based on 1) whether the idea expressed in the article is convincing; 2) whether the idea is actionable and 3) whether the idea is well-presented (https://seekingalpha.com/article/2134803-how-much-does-seeking-alpha-pay-its-contributors).

[8] According to the same article written by Seeking Alpha CEO Ali Hoffmann and posted on its website on 10th April 2014, Seeking Alpha pays authors who contribute articles exclusive to Seeking Alpha. The base payment is $10 each 1,000 page views. For high quality analysis of stocks that lack god research (e.g., small –cap), Seeking Alpha pays a minimum $150 for articles selected by its editors, and $500 for top small-cap ideas with exceptionally attractive risk/reward profiles.

[9] To the best of our knowledge, Seeking Alpha data has been used in only one published study. Chen *et al.* (2014) analyse articles published on Seeking Alpha between 2005 and 2012, and find that views expressed in

Secondly, our study provides new insights into the determinants of stock return co-movement. Since Roll (1988), a considerable amount of research has identified links between co-movement and investor protection and development of financial market, corporate governance, mandatory adoption of XBRL and newspaper coverage (Morck *et al.*, 2000; Durnev *et al.*, 2003; Crawford *et al.*, 2012; Dong and Ni, 2014; Dong *et al.*, 2016). Our study points to the role of social media in influencing stock return co-movement.

Finally, our results highlight that social media coverage has incremental effect on the incorporation of firm-specific information into stock price for firms with higher financial reporting opacity, which suggests that to a certain extent social media can complement formal disclosure (e.g., financial statements) in unravelling value-relevant information to the market. Our findings have implication for regulators, because regulators could encourage the use of social media as an alternative outlet of disclosing value-relevant information, and regulators may even coordinate with social media platforms to facilitate the dissemination of corporate information and an independent verification and evaluation of such information to a broader audience.

The remainder of the paper is organized as follows. Section 2 develops the model. Sample and research design are described in Section 3, and Section 4 presents the empirical results. The final section concludes.

2. The model

In the literature, security prices are considered to be jointly determined by a common market factor and an idiosyncratic factor (Roll, 1988; Jin and Myers, 2006).[10] When a larger fraction

articles and commentaries predict future stock returns and earnings surprises. Our research question is fundamentally different from Chen *et al.* (2014), in that we are interested in whether Seeking Alpha coverage facilitates the disclosure of firm-specific information to the market, thus enhancing overall price co-movement.

[10] Roll (1988) suggests that the idiosyncratic factor is largely determined by firm-specific information.

of price fluctuation is caused by the idiosyncratic factor, the co-movement between assets return and the market factor decreases. In contrast, if the market factor drives the price fluctuation to a greater extent, it is reasonable that asset returns will co-integrate with the market factor, generating higher return co-movement. Our two-period model, which is adapted from the model proposed by Huang *et al*., (2018), provides insights into how social media coverage and financial reporting opacity relate to the attention allocation of investors, which consequently influence the return co-movement. In this section, we briefly outline the key intuitions and conclusions of our model for conciseness, and refer the readers to Appendix 1 for a complete exposition of the model.

In our two-period model, we assume that the stock price movement is determined by a market shock, an industry shock and a firm-specific shock, and we assume one representative investor in the economy. The investor has no prior information about each shock at the start of the first period, but can learn and refine her beliefs about these shocks, which will be revealed at the end of the second period. The investor learn about these shocks by optimally assigning her limited amount of attention to each type of shock to obtain signals of the shocks. The signal for one type of shock is more precise if the investor assigns more attention to that particular type of shock. Consequently, conditioning on knowing the signals about these shocks, the investor is less uncertain about their final wealth in the end.

The investor maximizes her final utility by choosing the optimal demand of each security and optimal allocation of attention to each type of shock. Subsequently, the price of each asset at the end of the first period is determined based on the optimal demand conditioning on the signals. Return co-movements with market or industry returns for the first period are determined by the precision of the signal of firm-specific shocks. In detail, the more precise the signal of a firm-specific shock, the smaller return co-movements between market and industry returns. This results from the fact that the investor possesses more information about the firm-specific shock, so that the firm-specific shock contributes more to the stock price relative to the market and industry shocks.

Our contribution to the original model of Huang et al. (2018) is that we consider explicitly the information processing cost in the model. In our model, the information processing cost is directly associated with the attention allocated to the firm-specific shock. Intuitively, for the same amount of attention allocation, the firm-specific signal is less (more) accurate for the firm with a larger (smaller) information processing cost. We then link information processing cost to a firm's social media coverage and financial reporting opacity directly based on our hypothesis in the paper, such that firms with lower (higher) social media coverage and higher (lower) financial reporting opacity has a higher (lower) information processing cost. Based on this assumption, we derive two important propositions:

Proposition 1: Co-movement between firm-specific return and market/industry return is smaller for firms with a higher level of social media coverage.

Proposition 2: The marginal effect of social media coverage on return co-movement is more pronounced for firms with higher financial reporting opacity.

In the remainder of this paper, we test propositions 1 and 2 derived from our theoretical two-period model empirically using fixed-effect regression models. Details of the design of the tests can be found in Section 3.2.

3. Data, variable and research design

3.1 Data

Our measure of social media coverage is the number of articles exclusively related to a firm for a given year (single-ticker articles) that are posted on the Seeking Alpha, the largest social media platform specializing in financial analyses in the US. We design a computer program to automate the process of extracting all single-ticker articles from Seeking Alpha website. Specifically, we use a python program based on "Scrapy" to extract all single-ticker articles from the website in HTML format. Our data-set includes 133,217 single-ticker articles between 2004 and 2014. In the subsequent analysis we use the natural log of one plus the number of Seeking Alpha article to alleviate its skewness. Appendix 2 provides an example of a typical Seeking Alpha article. We further collect return data from CRSP, firm fundamental data from COMPUSTAT, and analyst coverage data from I/B/E/S. We require non-missing values for key variables including Seeking Alpha coverage, synchronicity and control variables listed in Section 3.2. Our final sample contains 39,568 firm-year observations from between 2004 and 2014.[11] To mitigate the potential influence of outliers, we winsorise all continuous variables at the top and bottom 1%.

3.2 Research design

3.2.1 Test of Seeking Alpha coverage and stock return co-movement

Roll (1988) is the first to propose that stock price synchronicity, the association between a firm's stock return and market and industry returns, is negatively associated with the amout of firm-specific information being impounded into stock prices. Consistent with this approach, we first estimate the adjusted R^2 from the following regression for each firm-year:

$$RET_{i,t} = \alpha_0 + \beta_1 \mathrm{Re}t\_mkt_{i,t} + \beta_2 \mathrm{Re}t\_ind_{i,t} + \beta_3 \mathrm{Re}t\_mkt_{i,t-1} + \beta_4 \mathrm{Re}t\_ind_{i,t-1} + \varepsilon_{i,t} \qquad (1)$$

where *RET* is the weekly stock return of individual firm *i* in week *t*; *Ret_mkt* is the weekly return calculated as the value-weighted return on all NYSE, AMEX, and NASDAQ stocks in week *t*; *Ret_ind* is the weekly return of the industry (based on the two-digit SIC code) in

[11] The number of observations used to test proposition 1 and 2 varies between 39,568 and 12,606, as the data requirements we impose to calculate CORRE and measures of financial reporting opacity result in loss of some of the observations. However, when we investigate proposition 1 and 2 with a reduced sample including the same number of observations, our inferences remain qualitatively unchanged.

week t to which the firm belongs. The lag returns are included to account for non-synchronous trading. Adjusted R^2 is derived from Equation 1. We run Regression (1) across each firm-year with a minimum of 45 weekly observations. Following previous literature (i.e., Morck *et al.*, 2000; Piotroski and Roulstone, 2004), we define synchronicity as:

$$SYNCH = \log\left(\frac{Adj.R^2}{1 - Adj.R^2}\right) \tag{2a}$$

The benefit of log transformation of $Adj.R^2$ is the creation of an unbounded variable out of a variable originally bounded between 0 and 1, which generates a dependent variable with approximately normal distribution. By construction, higher value of stock price synchronicity (*SYNCH*) indicates that firms' stock returns are closely tied to market and industry returns (higher return co-movement).

Our second measure of co-movement is the percentage point of time-series Pearson correlation coefficient between weekly firm return and weekly market return.

For time series return of firm i R_i and market return R_m, CORRE is computed as follows:

$$CORRE_{i,m} = \frac{COV(R_i, R_m)}{\sigma_{Ri}\sigma_{Rm}} \tag{2b}$$

To test the association between Seeking Alpha coverage and return co-movement, we estimate the following model:

$$Comovement_{i,t} = \alpha_0 + \beta_1 L\_SA_{i,t} + \beta_2 LNUM_{i,t} + \beta_3 SIZE_{i,t} + \beta_4 LMB_{i,t} + \beta_5 LEVERAGE_{i,t} + \beta_6 ROA_{i,t}$$
$$+ \beta_7 NIND_{i,t} + \beta_8 HERFSALE_{i,t} + \beta_9 STDROA_{i,t} + \beta_{10} BIG4_{i,t} + \sum \alpha_i Firm_i + \sum \alpha_j Year_j + \varepsilon_{i,t}$$
(3)

where *Comovement* is either stock price synchronicity or CORRE for firm *i* in year *t*, and $L\_SA_{i,t}$ is defined as the natural log of one plus the number of Seeking Alpha single-ticker articles covering firm *i* in year *t*. We also incorporate a set of control variables that have been identified in previous research as affecting stock price synchronicity (Roll, 1988; Hutton *et al.*, 2009; Crawford *et al.*, 2012; Kim and Shi, 2012). *LNUM* is constructed as the natural log of 1 plus the average number of analysts following the firm during the previous fiscal year. *SIZE* is defined as the natural log of the firm's market capitalisation at the end of the previous fiscal year; *LMB* is measured as the natural log of market capitalisation scaled by the book value of equity at the end of the previous fiscal year; *LEVERAGE* is the total long-term debt scaled by the total assets at the end of the previous fiscal year, and *ROA* is measured as income before extraordinary items divided by total assets at the end of the previous fiscal year. *NIND* is the natural log of the number of firms in the industry to which firm *i* belongs; *HERFSALE* is the sum of the squared terms of the proportion of a firm's revenue to total revenue in the industry at the end of the previous fiscal year; *STDROA* is the standard deviation of *ROA* in the previous five years. To capture the impact of the quality of external auditing, Big4 is set to 1 if the firm is audited by one of the Big 4 audit firms (PwC, Deloitte, Ernst & Young or KPMG), 0 otherwise. We adjust the standard error for heteroscedasticity, serial and cross-sectional correlation using a two-dimensional cluster at the firm and year level (Petersen, 2009). Finally, we include firm fixed effect and year fixed effect to address firm-specific and time-series trends of stock price synchronicity. As suggested by Dyreng *et al.*, (2010), the utilisation of firm-fixed effect forces the firm to act as its own control, and our test essentially concentrates on within-firm variation. The definition of all variables is summarized in Appendix 3.

3.2.2 The moderating effect of financial reporting opacity

To measure financial reporting opacity, we first follow Hutton *et al.,* (2009) to calculate the three-year average accruals quality over year *t-2* to year *t*, and label it *OPA*. Discretionary accruals are computed as residual from the estimation error Model (Equation 4). Following Francis *et al.*, (2005) model, we calculate the absolute value of the residual from each year for the two-digit SIC industry, and larger discretionary accruals (larger absolute value of residual) indicate lower accrual quality:

$$\frac{TACC_{i,t}}{TA_{i,t-1}} = \lambda_0 + \frac{\lambda_1 CFO_{i,t-1}}{TA_{i,t-1}} + \frac{\lambda_2 CFO_{i,t}}{TA_{i,t-1}} + \frac{\lambda_3 CFO_{i,t+1}}{TA_{i,t-1}} + \frac{\lambda_4 \Delta \mathrm{Re}v_{i,t}}{TA_{i,t-1}} + \frac{\lambda_3 PPE_{i,t}}{TA_{i,t-1}} + \varepsilon_{i,t} \quad (4)$$

Our second measure of financial reporting opacity is the dispersion of analyst earnings forecasts. Previous research (Maffett, 2012) shows that accounting information asymmetry is associated with greater analyst forecast dispersion, as the disagreement of earnings forecasts among analysts reflects the level of information asymmetry between corporate insiders and outsiders (i.e. investors and intermediaries such as analysts). Opaque firms disclose less firm-specific information or information of inferior quality, which makes it difficult for analysts to reach a consensus on earnings forecasts, generating large forecast dispersion. We compute the dispersion of analyst earnings forecasts for each firm in year *t*, and create *OPA2* by scaling the dispersion with the firm's opening stock price of the year.

Then we rely on the following model to test whether Seeking Alpha coverage plays a more pronounced role in decreasing return co-movement among firms with higher financial reporting opacity:

$$\begin{aligned} Comovement_{i,t} &= \alpha_0 + \beta_1 L\_SA_{i,t} + \beta_2 OPA_{i,t} + \beta_3 L\_SA_{i,t} * OPA_{i,t} + \beta_4 LNUM_{i,t} + \beta_5 SIZE_{i,t} + \beta_6 LMB_{i,t} + \beta_7 LEVERAGE_{i,t} \\ &+ \beta_8 ROA_{i,t} + \beta_9 NIND_{i,t} + \beta_{10} HERFSALE_{i,t} + \beta_{11} STDROA_{i,t} + \beta_{12} BIG4_{i,t} + \sum \alpha_i Firm_i + \sum \alpha_j Year_j + \varepsilon_{i,t} \end{aligned}$$

(5)

where *Comovement* is either stock price synchronicity or CORRE for firm *i* in year *t*. We run Regression (5) for the full sample, and expect the coefficient of interaction between Seeking Alpha coverage and OPA (OPA2), our main variable of interest, to be significantly negative.

4. Empirical results

4.1 Descriptive statistics

Figure 1 presents the summary statistics on the Seeking Alpha coverage. The number of Seeking Alpha articles (firms covered by Seeking Alpha) increased from 27 (19) in 2004 to

24,939 (3,969) in 2014, showing the substantially growing influence of Seeking Alpha in the investment community during our sample period.

Table 1 provides the descriptive statistics of all the variables. The synchronicity measure has a mean (median) of -2.664 (-2.421), and varies from -3.874 (25th percentile) to -1.234 (75th percentile). CORRE has a mean (median) of 40.898 (43.003) and varies from 24.361 (25th percentile) to 58.662 (75th percentile). The mean (median) of the number of Seeking Alpha articles (L_SA) is 0.426 (0.000). The mean (median) of analyst following is 0.952 (0.000), and mean (median) of firm size measured by the logarithm of market capitalisation is 6.213 (6.184), which suggests our sample is populated with large firms. ROA has a mean (median) of 0.025 (0.024), and standard deviation of ROA has a mean (median) of 0.101 (0.033). Finally, the mean (median) of Big4 is 0.701 (1.000), indicating that more than 70% of our sample firms have Big4 as their auditor. All the variables have substantial variation.

<< Insert Table 1 about here >>

4.2 Correlation

Table 2 presents the Pearson correlation between the variables. Both CORRE and the synchronicity measure are positively correlated with Seeking Alpha coverage, which seems inconsistent with our prediction.[12] However, as we do not control for the other determinants of the synchronicity measure, the correlation has to be interpreted with caution. Consistent with findings of Piotroski and Roulstone (2004), the two co-movement measures are positively correlated with analyst following, suggesting that high analyst coverage facilitates the disclosure of market and industry information to the market. The two co-movement

[12] The positive correlation likely captures the relation between synchronicity measure and firm size (correlation = 0.645, $p < 0.001$), which in turn is positively associated with Seeking Alpha coverage.

measures are positively correlated with size, market-to-book, leverage, ROA, Big4, and is negatively related to the standard deviation of ROA and *NIND*. Finally, the correlation statistics do not raise concerns regarding multicollinearity, as the largest correlation is that between CORRE and synchronicity (0.847) and they will not enter the same regression. The VIF of all subsequent regressions are below 10.

<< Insert Table 2 about here >>

4.3 Multivariate results

Table 3 reports results related to the prediction that Seeking Alpha coverage is associated with lower return co-movement. We use the logarithm of one plus the number of articles on Seeking Alpha (L_SA) as the proxy of Seeking Alpha coverage, and synchronicity (column 1 and 2) and CORRE (column 3 and 4) are used as proxy of co-movement. In column 1(3) we regress synchronicity (CORRE) on Seeking Alpha coverage as well as analyst coverage and firm size as control variable, because the correlations between these two variables and synchronicity measure are the highest among the correlations between all control variables and synchronicity measure. In column 2(4) we employ the complete set of control variables in the analysis, and also control for both firm-fixed and time-fixed effects. As the results across the columns are consistent, we focus on the results in column 2 and 4 for interpretation. When synchronicity and CORRE are the dependent variable, the coefficient of L_SA is negative and significant (-0.059, *t = -4.137;* -0.493*; t= -2.595* respectively), which indicates that more Seeking Alpha coverage enables the incorporation of firm-specific information into stock price to a greater extent, leading to lower co-movement. The results lend credence to the contention that investment related information disclosed on social media can effectively be revealed to the market and capitalised into stock price. Furthermore, we calculate the marginal effect to gauge the economic significance. With all other variables at their sample mean, a standard deviation increase of L_SA from its mean results in a 4.3% (9.50%) reduction in stock price synchronicity (CORRE). Therefore we conclude that the negative association between Seeking Alpha coverage and return co-movement is both statistically and economically significant. We further decompose the synchronicity measure into a market return component and an industry return component, and in untabulated analysis find that Seeking Alpha coverage is negatively associated with both components with statistical significance.

In respect of the control variables, the coefficients of analyst coverage and size are significantly positive, whereas the coefficients of market-to-book are negative and significant. These findings are broadly consistent with those documented in previous research (Piotroski and Roulstone, 2004; Crawford *et al.*, 2012).

<< Insert Table 3 about here >>

Table 4 shows results related to the prediction that the influence of Seeking Alpha coverage on return co-movement is more pronounced in firms with high financial reporting opacity. We use 1) the three-year average of discretionary accruals calculated from the Francis *et al.*, (2005) model (OPA) 2) analyst forecast dispersion (OPA2) as proxies for financial reporting opacity, so firms with large average discretionary accruals (larger forecast dispersion) are considered to be more opaque. We introduce an interaction term between OPA (OPA2) and Seeking Alpha coverage measures (L_SA), and the coefficient of the interaction term is the main variable of interest.

We present 4 models, where OPA (OPA2) is the proxy of opacity in module 1 and 2 (3 and 4). In column 1 and 2 where synchronicity and CORRE are the dependent variable, the coefficient of the interaction between L_SA and OPA is negative and significant (-0.143, *t = -4.294;* -1.939, *t= -3.637* respectively), In column 3 and 4 where synchronicity and CORRE are the dependent variable, the coefficient of interaction between L_SA and OPA2 is significantly negative (-0.071, *t = -4.561;* -0.435, *t= -2.453* respectively). The findings suggest that relative to firms with lower opacity (smaller discretionary accruals and smaller forecast dispersion), Seeking Alpha coverage plays a more significant role in reducing return co-movement for firms with high opacity (larger discretionary accruals and larger forecast dispersion). It is likely that investors (in particular individual investors) find it difficult and costly to acquire information for firms with high opacity, and Seeking Alpha articles effectively enable the flow of firm-specific information on such firms to the market, leading to an incremental decrease in return co-movement.[13]

<< Insert Table 4 about here >>

4.4 Endogeneity issue

4.4.1 Propensity score matching (PSM)

Our results may suffer from endogeneity, because firms are less likely to be randomly covered by social media such as Seeking Alpha. For example, large firms, firms with extreme unexpected earnings or firms in selected industries (i.e. consumer-oriented industries) are more likely to be covered on Seeking Alpha. Our first approach to address the concern of endogeneity is the propensity score matching (PSM) method. We estimate the following logit model for each year: the dependent variable is coded 1 if a firm is covered by Seeking Alpha

[13] We divide our sample into low opacity and high opacity sub-samples based on a sample mean of discretionary accruals and analyst forecast dispersion, and run the baseline model in two sub-samples. Untabulated results show that both the magnitude and significance levels of the Seeking Alpha coverage measure are higher in the high-opacity sub-sample, suggesting that Seeking Alpha coverage plays a more significant role in reducing return co-movement among firms with high opacity. This is consistent with the findings based on regression analysis incorporating interaction between the measure of opacity and Seeking Alpha coverage.

in a given year and zero otherwise; the independent variables include all firm-level control variables in Equation 3. Secondly, without replacement we match each "treatment firm" (a firm covered by Seeking Alpha in a given year *t*) with two matching firms (firms that are not covered by Seeking Alpha in the same year) that have the closet propensity scores within a maximum distance of 1%. That is, we use a nearest-neighbour matching approach with common support and a caliper constraint of 0.01. We have 21,528 observations for this analysis. The matching appears successful as the standardized biases of variables are less than 5% after the matching. We include year fixed effect in the first stage prediction model, which effectively removes the time-trend of increasing Seeking Alpha coverage over the sample period.

We repeat the analysis using the PSM sample, and the results are reported in Table 5. It is clear that the tenor of our results remains qualitatively unchanged, because the coefficient of Seeking Alpha coverage is significantly negative (-0.055, *t = -3.138;* -1.133, *t= -5.111*) in regressions when synchronicity and CORRE are used as measure of co-movement.

<< Insert Table 5 about here >>

4.4.2 Two-stage least square (2SLS)

We employ two instrumental variables (IVs) for Seeking Alpha coverage to further mitigate the issue of endogeneity. The first instrument is the annual advertising expenditure of a firm. We construct the instrumental variable L_ADX using the natural log of 1 plus the firm's annual advertising expenditure. The second instrument is the intensity of Seeking Alpha coverage for the industry to which a firm belongs. The second instrumental variable, denoted by L_SA_IND is measured as the natural log of 1 plus the total annual number of Seeking Alpha articles for the industry to which the firm belongs. We expect Seeking Alpha coverage to be positively correlated with both IVs, because firms with higher advertising expenditure are more likely to attract the attention of both investors in general and registered users of Seeking Alpha in particular. Firms belonging to industries that attract more attention from Seeking Alpha are more likely to be covered. On the other hand, firm-level advertising expenditure and industry-level Seeking Alpha coverage are less likely to have a direct influence on return co-movement at firm level. Both IVs pass the over-identification test, and the results are consistent with our prediction. In the first stage, we find that both IVs are significantly and positively associated with Seeking Alpha coverage at firm level. In the second stage, the negative and significant effect of predicted Seeking Alpha coverage on co-movement remains, thus corroborating our findings reported in Section 4.3.

<< Insert Table 6 about here >>

4.4.3 Loss of social media coverage

We further address the endogeneity issue by concentrating on the loss of Seeking Alpha coverage. Seeking Alpha authors who persistently publish articles and whose articles receive more commentaries would attract more attention from the audience. Therefore, in each quarter we rank all contributing authors based on their number of articles, number of firms covered, total comments received and average comments received per article. The quarterly synchronicity measure is calculated based on the adjusted R^2 from regressions of daily stock returns against the market return and the industry returns for each firm-quarter observation:

$$RET_{i,t} = \alpha_0 + \beta_1 \mathrm{Re}t\_mkt_{i,t} + \beta_2 \mathrm{Re}t\_ind_{i,t} + \varepsilon_{i,t} \tag{6}$$

We also calculate CORRE as the time-series Pearson correlation coefficient between daily firm returns and the market returns for each quarter.

Loss of Seeking Alpha coverage is defined when an author who meets the following criteria stops publishing articles on Seeking Alpha: 1) the author must have been publishing for at least four quarters continuously; 2) during the continuous coverage period the author's rank in terms of quarterly number of articles published, number of firms covered, total comments received and average comments received per article stay in the top 50% of all contributing authors.[14] We expect that the loss of coverage by such an influential author results in less firm-specific information flow to the market, leading to stock price being more synchronous to the market (higher return co-movement). For each case of loss of coverage, we calculate the quarterly average coverage of the firm that the author published during her continuous-publishing period. L_SA_DROP is defined as the natural log of 1 plus the average number of Seeking Alpha articles for a firm contributed by the author before she ceases publishing articles in the following quarter. Using the specified criteria, we find 366 exogenous drop events of active authors, which leads to 6,379 exogenous drops of firm-quarter observations. We run Model 7 to examine the potential effect of loss of coverage by active SA authors. The higher the value of L_SA_DROP, the more severe in increase in return co-movement after the loss of coverage, as evidenced by a significant and positive β_2.

$$\begin{aligned}&Comovement_{i,t} = \alpha_0 + \beta_1 L\_SA_{i,t} + \beta_2 L\_SA\_WDROP_{i,t-1} + \beta_3 LNUM_{i,t} + \beta_4 SIZE_{i,t}\\&+\beta_5 LMB_{i,t} + \beta_6 LEVERAGE_{i,t} + \beta_7 ROA_{i,t} + \beta_8 NIND_{i,t} + \beta_9 HERFSALE_{i,t} + \beta_{10} STDROA_{i,t} + \beta_{11} BIG4_{i,t}\\&+\sum \alpha_i Firm_i + \sum \alpha_j Quarter_j + \varepsilon_{i,t}\end{aligned} \tag{7}$$

The results, which are presented in Table 7, Column 2 (where the dependent variable is synchronicity) and column 4 (where the dependent variable is CORRE), are consistent with our prediction (β_2 = 0.162, *t = 2.033;* β_2 = 1.875, *t = 1.871* respectively). The findings show that loss of coverage indeed results in higher co-movement for the previously covered firm,

[14] The median of our active author sample has an average quarterly coverage of three firms. An active author may discontinue publishing articles related to a specific firm due to the dynamic of company-specific information. However, loss of entire coverage by an active author is probably due to personal reasons.

which implies a casual relationship between Seeking Alpha coverage and return co-movement.

<< Insert Table 7 about here >>

4.4.4 "Day of the Week" effect

We test the "day of the week effect" based on the conjecture that articles published on weekdays (Monday-Thursday) are more influential, as investors can trade in response to the news released in an article without delay. In contrast, investors have to wait for the trading day in the following week before they can react to information released in articles published during weekends. Therefore, our prediction is that articles published on weekend (Friday, Saturday and Sunday) have a relatively smaller information effect. We test such a conjecture with Model 8, with the expectation that β_1 (the coefficient of weekend coverage) is significantly smaller than β_2 (the coefficient of weekday coverage).

$$\begin{aligned} & Comovement_{i,t} = \alpha_0 + \beta_1 L\_SA\_WEEKENDS_{i,t} + \beta_2 L\_SA\_WEEKDAY_{i,t} + \beta_3 LNUM_{i,t} + \beta_4 SIZE_{i,t} \\ & + \beta_5 LMB_{i,t} + \beta_6 LEVERAGE_{i,t} + \beta_7 ROA_{i,t} + \beta_8 NIND_{i,t} + \beta_9 HERFSALE_{i,t} + \beta_{10} STDROA_{i,t} + \beta_{11} BIG4_{i,t} \\ & + \sum \alpha_i Firm_i + \sum \alpha_j Quarter_j + \varepsilon_{i,t} \end{aligned} \quad (8)$$

L_SA_WEEKENDS is the natural log of 1 plus the number of single-ticker Seeking Alpha articles posted on Friday, Saturday and Sunday for a firm in the quarter. L_SA_WEEKDAY is the natural log of 1 plus the number of single-ticker Seeking Alpha articles posted from Monday to Thursday for a firm in the quarter. The results, which are presented in Table 8, are consistent with our prediction (β_1 = -0.079, t = *-3.529;* β_2 = -0.167, *t* = *-9.963* when synchronicity is the dependent variable; β_1 = -1.204, t = *-5.397;* β_2 = -2.048, *t* = *-12.239* when CORRE is the dependent variable. F-test confirms that in both regressions β_1 is significantly smaller than β_2).

<< Insert Table 8 about here >>

4.5 Robustness check

In this section we test whether Seeking Alpha coverage facilitates the incorporation of a type of firm-specific information, future earnings, into current stock price. We use the model outlined in Kothari and Sloan (1992):

$$RET_{i,t,t-k} = \rho_0 + \rho_1 * E_{i,t} + \varepsilon_{i,t} \tag{9}$$

where $RET_{i,t,t-k}$ is the stock return from period *t–k* to period *t*. $E_{i,t}$ is the earnings in period *t* (future earnings), defined as income before extraordinary items divided by total assets. As predicted by Kothari and Sloan (1992), when the time interval *k* increases, firm-level information becomes more likely to be incorporated into the return over the period *t–k* to *t*. Hence, ρ_1 should increase with *k*. In the case where investors retrieve firm-level information earlier, the estimated ρ_1 will be larger when the estimated interval is longer. We estimate the following system of equations using the Seemingly Unrelated Regression (SUR) technique, as SUR provides more efficient estimation than separate OLS regressions when the disturbances of the equations are related.

$$RET_{i,t,t-1} = \rho_0 + \rho_1 EARNING_{i,t} + \rho_2 EARNING_{i,t} * L\_SA_{i,t-1} + Controls + \varepsilon_{i,t} \qquad 9(a)$$

$$RET_{i,t,t-2} = \theta_0 + \theta_1 EARNING_{i,t} + \theta_2 EARNING_{i,t} * L\_SA_{i,t-2} + Controls + \varepsilon_{i,t} \qquad 9(b)$$

The ratio of θ_2 in Equation 9(b) to ρ_2 in Equation 9(a) measures the relative speed with which stock price incorporates future earnings for firms with higher Seeking Alpha coverage. The larger the ratio, the earlier investors incorporate a firm’s future earnings into current stock price for firms with higher Seeking Alpha coverage. Evidence supporting such prediction is consistent with the notion that higher Seeking Alpha coverage is associated with more forward-looking information being capitalised into current stock price.

Table 9 presents results supporting the view that Seeking Alpha coverage facilitates the incorporation of future earnings into current stock price. In particular, θ_2 in the two-year period estimation (0.056, *t = 6.972*) is significantly larger than ρ_2 in the one-year estimation (0.012, *t = 3.342*). The difference of coefficients results in a Wald-statistics of 23.07 (p < 0.01). To conclude, we find evidence that stock prices of firms with higher Seeking Alpha coverage incorporate future earnings more efficiently.

<< Insert Table 9 about here >>

Our final robustness check utilizes an exogenous shock launched by Seeking Alpha in 2011. From January 2011, Seeking Alpha started paying each contributing authors $10 per 1,000 page views on their articles, with an attempt to attract submission with high quality of information. Therefore, we expect that in the post-2011 period Seeking Alpha articles provided more information to the market relative to that in the pre-2011 period. Consequently, the impact of Seeking Alpha coverage on co-movement is likely to be more significant in the post-2011 period.[15] To test this prediction, we introduced an interaction term between L_SA and a dummy variable POST (which takes 1 for observations after 2011 and 0 otherwise). We maintain the same set of control variables as in early analysis.[16] The results, which are presented in Table 10, lend support to our conjecture. The coefficient of interaction is negative and significant when synchronicity and CORRE are the dependent variable (-0.140, t= *-3.720*; -3.652, *t= -3.056*), suggesting that Seeking Alpha coverage plays

[15] The event year 2011 is not considered as a post-2011 year to avoid potential confounding effect. Therefore, the post-2011 period includes 2012, 2013 and 2014.

[16] Because POST is a time dummy, we do not include time fixed effect in the analysis. Instead we cluster the standard error by firm and year to account for autocorrelation.

a stronger role in reducing co-movement after financial incentive has been provided to enhance the information quality of published articles in 2011.

<< Insert Table 10 about here >>

We conducted additional robustness checks. For brevity we do not tabulate these results. As Seeking Alpha articles that receive more commentaries would attract more attention from the audience, we classified articles into influential ones and less influential ones based on the median value of commentaries on the original article for a given year, and expected influential articles to play a more significant role in reducing return co-movement. The untabulated results confirm our prediction that influential Seeking Alpha articles have more effect on reducing return co-movement than less influential articles.

5. Conclusion

Mindful of the increasing importance of social media as a venue for information production and dissemination in the new millennium, in this paper we develop a model to predict that the coverage of a public firm on social media substantially reduces the information acquisition and processing cost for investors. Consequently, this facilitates the transmission of more firm-specific information into stock price, resulting in lower return co-movement. We test the prediction of our model with data collected from the Seeking Alpha website between 2004 and 2014, and find that Seeking Alpha coverage is negatively associated with return co-movement. In addition, we show that the effect of Seeking Alpha coverage on return co-movement is more salient in firms with higher financial reporting opacity. Our findings are robust to propensity score matching, a two-stage least square (2SLS) approach, and alternative measures of firm opacity.

Our study is of interest to investors because sophisticated financial market participants might be incentivized to develop trading techniques that take into account the coverage on social media when formulating their trading strategy. Our findings have important implication for regulators, as social media is landscape shifting in that it has become a revolutionary approach to information generation, evaluation and dissemination in the 21st century due to its global access and interactive nature.

References

Barberis, N., A. Shleifer, and J. Wurgler. (2005). Co-movement. *Journal of Financial Economics*, *75*(2), 283–317.

Beuselinck, C., P. Joos, I. Khurana, and S. Van der Meulen. (2010). Mandatory IFRS reporting and stock price informativeness. Working paper, Tilburg University.

Bikhchandani, S., D. Hirshleifer, and I. Welch. (1998). Learning from the behavior of others: Conformity, fads, and informational cascades. *Journal of Economic Perspectives*, *12*, 151–170.

Blankespoor, E., G. Miller, and H. White. (2014). The role of dissemination in market liquidity: Evidence from firms' use of Twitter. *Accounting Review*, *89*, 79–112.

Bleck, A., and X. Liu. (2007). Market transparency and the accounting regime. *Journal of Accounting Research*, *45*, 229–256.

Bushman, M., and J. Smith. (2001). Financial accounting information and corporate governance. *Journal of Accounting and Economics*, *32*(1), 237–333.

Chen, H, P. De, Y. Hu, and B. Hwang. (2014). Wisdom of crowds: The value of stock opinions transmitted through social media. *Review of Financial Studies*, *27*, 1367–1403.

Cohen, L., and D. Lou. (2012). Complicated firms. *Journal of Financial Economics*, *104*, 383–400.

Collins, D. W., S. P. Kothari, J. Shanken, and R. Sloan. (1994). Lack of timeliness and noise as explanation for the low contemporaneous return–earnings association. *Journal of Accounting and Economics*, *18*, 289–324.

Crawford, S., D. Roulstone, and E. So. (2012). Analyst initiations of coverage and stock price efficiency. *Accounting Review*, *87*, 1527–1553.

Dechow, M., G. Sloan, and P. Sweeney. (1995). Detecting earnings management. *Accounting Review*, *70*, 193–225.

Dong, Y., and C. Ni. (2014). Does limited attention constrain investors' acquisition of firm-specific information? *Journal of Business Finance and Accounting*, *41*, 1361–1392.

Dong, Y., O. Z. Li, Y. Lin, and C. Ni. (2016). Does information processing cost affect firm-specific information acquisition? Evidence from XBRL adoption. *Journal of Financial and Quantitative Analysis*, *51*, 435–462.

Drake, M., J. Thornock, and B. Twedt. (2017). The Internet as an information intermediary. *Review of Accounting Studies*, *22*,1-34.

Durnev, A., R. Morck, B. Yeung, and P. Zarowin. (2003). Does greater firm-specific return variation mean more or less informed stock pricing? *Journal of Accounting Research*, *41*, 791–836.

Dyreng, S., M. Hanlon, and E. Maydew, 2010. The effects of executives on corporate tax avoidance. *Accounting Review*, *85*, 1163–1189.

Fernandes, N. and M. Ferreira. (2008). Does international cross-listing improve the information environment? *Journal of Financial Economics*, *88*, 216–244.

Fernandes, N., and M. Ferreira. (2009). Insider trading laws and stock price informativeness. *Review of Financial Studies*, *22*, 1845–1887.

Ferreira, M., and P. Laux. (2007). Corporate governance, idiosyncratic risk, and information flow. *Journal of Finance*, *62*, 951–989.

Francis, J., R. LaFond, P. Olsson, & K, Schipper. (2005). The market pricing of accruals quality. Journal of accounting and economics, 39(2), 295-327.

Grossman, S., and J. Stiglitz. (1980). On the impossibility of informationally efficient markets. *American Economic Review*, *70*, 393–408.

Gul, F. A., B. Srinidhi, and A. C. Ng. (2011). Does board gender diversity improve the informativeness of stock prices? *Journal of Accounting and Economics*, *51*, 314–338.

Haggard, S., X. Martin, and R. Pereira. (2008). Does voluntary disclosure improve stock price informativeness? *Financial Management*, *37*, 747–768.

Haw, I., B. Hu, J. J. Lee, and W. Wu. (2012). Investor protection and price informativeness about future earnings: International evidence. *Review of Accounting Studies*, *17*, 389–419.

Hirshleifer, D., Teoh, S. (2003). Limited attention, information disclosure and financial reporting. *Journal of Accounting and Economics*, 35, 337-386.

Hirshleifer, D., Lim, S., Teoh, S. (2003). Disclosure to a credulous audience: the role of limited attention. *Working paper*, The Ohio State University.

Hope, O. (2003). Disclosure practices, enforcement of accounting standards, and analysts' forecast accuracy: An international study. *Journal of Accounting Research*, *41*, 235–272.

Huang, S., Y. Huang, and T.-C. Lin (2018). Attention allocation and return co-Movement: Evidence from repeated natural experiments. *Journal of Financial Economics*, Forthcoming.

Hutton, P., A. Marcus, and H. Tehranian. (2009). Opaque financial reports, R^2, and crash risk. *Journal of Financial Economics*, *94*(1), 67–86.

Ivković, Z., and S, Weisbenner. (2007). Information diffusion effects in individual investors' common stock purchases: Covet the neighbors' investment choices. *Review of Financial Studies*, *20*, 1327–1357.

Jin, L., and S. Myers. (2006) R^2 around the world: New theory and new tests. *Journal of Financial Economics*, *79*, 257–292.

Kim, J., and H. Shi. (2012). IFRS reporting, firm-specific information flows, and institutional environments: International evidence. *Review of Accounting Studies*, *17*, 474–517.

Kothari, S. P., and R. Sloan. (1992). Information in prices about future earnings. *Journal of Accounting and Economics*, *15*, 143–171.

Lambert, A. (2001). Contracting theory and accounting. *Journal of Accounting and Economics*, *32*, 3–87.

Lee, F., A. Hutton, and S. Shu. (2015). The role of social media in the capital market: Evidence from consumer product recalls. *Journal of Accounting Research, 53*, 367–404.

Li, K., R. Morck, F. Yang, and B. Yeung. (2004). Firm-specific variation and openness in emerging markets. *Review of Economics and Statistics*, *86*, 658–669.

Maffett, M. (2012). Financial reporting opacity and informed trading by international institutional investors. *Journal of Accounting and Economics*, *54*, 201–220.

Morck, R., B. Yeung, and W. Yu. (2000). The information content of stock markets: Why do emerging markets have synchronous stock price movements? *Journal of Financial Economics*, *58*, 215–260.

Petersen, M. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies*, *22*, 435–480.

Piotroski, J., and D. Roulstone. (2004). The influence of analysts, institutional investors and insiders on the incorporation of market, industry and firm-specific information into stock prices. *Accounting Review*, *79*, 1119–1151.

Roll, R. (1988). R^2. *Journal of Finance*, *43*, 541–566.

Shiller, R. (1995). Conversation, information and herd behavior. *American Economic Review*, *85*, 181–185.

Shiller, R. (1999). Human behavior and the efficiency of the financial system. In J. B. Taylor and M. Woodford (eds.), *Handbook of Macroeconomics*, *1*, 1305–1340.

Shiller, R., and J. Pound. (1989). Survey evidence on diffusion of interest and information among investors. *Journal of Economic Behavior and Organization*, *12*, 47–66.

Tucker, J. W., and P. Zarowin. (2006). Does income smoothing improve earnings informativeness? *Accounting Review*, *81*, 251–270.

Yu, J. (2011). Stock price informativeness and corporate governance: An international study. *International Review of Finance*, *11*, 477–514.

Appendix 1: The two-period model

Our two-period model closely follows the model proposed by Huang *et al*., (2018) based on the theoretical framework of Kacperczyk *et al.*, (2016). We provide insights on how social media coverage and financial reporting opacity relate to the attention allocation of attention of investors, which further influence the co-movement between firm specific returns and the industry and market returns.

1.1 General setting and asset payoffs

In the model we consider an economy of two periods: $t = 0,1,2$. At time 0, investors have no prior beliefs about the payoffs of the securities. At time 1, investors update their beliefs by allocating their attentions to different securities, and rebalance their portfolio. The payoffs of the assets are realized at time 2.

We assume there are n>2 industries and $k > 2$ securities within each industry. Each stock has a net positive supply normalized as 1. There exists a risk-free asset with the payoff of 1 for each unit. The payoff of the j-th stock in the i-th industry is denoted as $V_{i,j}$, which is realized at $t = 2$. The following relationship is assumed for the payoff of the j-th stock in the i-th industry for all $j = 1:k$ and $i = 1:n$:

$$V_{i,j} = \bar{V}_{i,j} + m + e_i + f_{i,j},$$

in which $\bar{V}_{i,j}$ is the mean payoff of the security, m is the market-wide shock to the payoff, e_i is an industry-wide shock to the payoff, and $f_{i,j}$ is a firm-specific shock to the payoff. Investors have no information about these shocks at time 0. We assume the following dynamics of the shocks:

$$m \sim \mathrm{N}\,(0, \frac{1}{\tau_m}),$$

$$e_i \sim \mathrm{N}\,(0, \frac{1}{\tau_e}),$$

$$f_{i,j} \sim \mathrm{N}\,(0, \frac{1}{\tau_f}),$$

$$m, e_i, f_{i,j} \text{ mutually independent.}$$

Following Huang *et al.*, (2018) and to simplify the analysis, we assume that $\tau_m = \tau_e = \tau_f = \tau > 0$, so that the variance of three shocks are identical.

1.2 Investors and learning

There is one representative investor in the economy, and we assume that the investor has a mean-variance utility function over his wealth at $t = 2$. For each $t = 0,1,2$, we denote his wealth at time t as W_t. We use F_t to denote the information set available to the investor at time t. Clearly we have $\mathrm{F}_0 \subset \mathrm{F}_1 \subset \mathrm{F}_2$, where F_0 is the trivial information set at time 0. At time 1, investor updates their information set to F_1 by refining their beliefs about the shocks m, e_i and $f_{i,j}$. The information set F_2 contains all the information about the payoffs, and is not available to the investor. The investor's objective is to optimize the following utility function based on the available information set:

$$\mathrm{E}(W_2 \mid \mathrm{F}_1) - \frac{\gamma}{2}\mathrm{V}(W_2 \mid \mathrm{F}_1),$$

where γ is understood as the risk aversion coefficient.

At time 1, the investor can learn about the shocks m, e_i and $f_{i,j}$ by allocating her attention, which is constrained by some positive constant K, known as the learning capacity. We propose to specify the learning process of the investor as follows:

$$
\begin{aligned}
&s_m = m + \varepsilon_m, \quad \varepsilon_m \sim \mathrm{N}\,(0, \frac{1}{\kappa_m}), \\
&s_{e,i} = e_i + \varepsilon_{e,i}, \quad \varepsilon_{e,i} \sim \mathrm{N}\,(0, \frac{1}{\kappa_{e,i}}), \\
&s_{f,i,j} = f_{i,j} + \varepsilon_{f,i,j}, \quad \varepsilon_{f,i,j} \sim \mathrm{N}\,(0, \frac{1}{\kappa_{f,i,j} E_{i,j}}), \\
&E_{i,j} = \alpha_{i,j} + O_{i,j}, \\
&\varepsilon_m, \varepsilon_{e,i}, \varepsilon_{f,i,j} \text{ mutually independent,}
\end{aligned}
$$

in which s_m, $s_{e,i}$ and $s_{f,i,j}$ are the signals the investor receive at time 1 about the market shock m, the industry shock e_i and the firm-specific shock $f_{i,j}$. For the market and industry signals, they are obtained with errors ε_m and $\varepsilon_{e,i}$ respectively. The investor needs to allocate her attention κ_m and $\kappa_{e,i}$ to each error term to refine the signal. Intuitively, larger κ_m results in a more precise signal for the market component, and as $\kappa_m \to 0$, s_m is not informative about m at all.

Deviating from Huang *et al*., (2018), we propose to introduce additional quantities $\alpha_{i,j}$, and $O_{i,j}$ for each stock to refine investor's learning process about the firm-specific shock $f_{i,j}$. The quantity $\alpha_{i,j} \geq 0$ stands for the social media coverage for the stock at time 1, such that larger $\alpha_{i,j}$ represents a higher social media coverage, and vice versa. The quantity $O_{i,j} \geq 0$ is defined as a measure of the transparency of the stock's financial reports, that is, the larger $O_{i,j}$ is, the less opaque the firm's financial reports. In our model we assume that $\alpha_{i,j}$ and $O_{i,j}$ are independent across firms.

In the specification above, for the same amount of attention input $\kappa_{f,i,j}$, larger $\alpha_{i,j}$ and $O_{i,j}$ reduce the information processing cost according to our hypothesis, and vice versa. By multiplying the attention allocation $\kappa_{f,i,j}$ by $E_{i,j} = \alpha_{i,j} + O_{i,j}$, we model the information processing cost of learning the firm specific shock explicitly such that any increase in $\alpha_{i,j}$ or

$O_{i,j}$ decreases the information cost and improves the precision of firm-specific signal $s_{f,i,j}$ for a given attention allocation $\kappa_{f,i,j}$. The quantity $E_{i,j}$ can be interpreted as an inverse measure of information processing cost, and clearly when $E_{i,j}=1$ for all i and j, our model reduces to the model in Huang *et al*., (2018).

At time 1, the investor will decide how to allocate her attentions subject to her total learning capacity K:

$$\kappa_m + \sum_{i=1}^{n} \kappa_{e,i} + \sum_{i=1}^{n} \sum_{j=1}^{k} \kappa_{f,i,j} \leq K.$$

We also require that κ_m, $\kappa_{e,i}$ and $\kappa_{f,i,j}$ are larger or equal to zero, so the investor cannot allocate negative attention.

1.3 Price, demand and optimal attention allocation at equilibrium

We denote the nk-by-1 vector of asset prices at time t by $\mathbf{P}_t$, which is of the following structure:

.
$$\mathbf{P}_t = \{P_{i,j,t}\}_{i=1:n,\, j=1:k},$$

where $P_{i,j,t}$ is the price of the j-th stock in the i-th industry at time t. It is obvious that $\mathbf{P}_2 = \mathbf{V}$, where $\mathbf{V}$ is the vector of payoff from all stocks at time 2. Similarly we define the nk-by-1 vector of asset prices at time t as:

$$\mathbf{X}_t = \{X_{i,j,t}\}_{i=1:n,\, j=1:k},$$

where $X_{i,j,t}$ is the demand of the j-th stock in the i-th industry at time t. The wealth of the investor can therefore be expressed as:

$$W_2 = W_0 + \mathbf{X}_0^T(\mathbf{P}_1 - \mathbf{P}_0) + \mathbf{X}_0^T(\mathbf{V} - \mathbf{P}_1).$$

The investor maximizes her utility at time 2 through the following two-step procedure. At $t=0$, the investor does not have any meaningful information about the shocks, and will maximize her utility function conditional on the trivial information set F_0, which is omitted for brevity:

$$\max_{\mathbf{X}_0} \mathrm{E}[W_2] - \frac{\gamma}{2}\mathrm{V}[W_2],$$
$$s.t. \quad W_2 = W_0 + \mathbf{X}_0^T(\mathbf{P}_1 - \mathbf{P}_0) + \mathbf{X}_1^T(\mathbf{V} - \mathbf{P}_1).$$

At time 1, the investor obtain new information set F_1, and need to maximize her utility based on the new information set through the following two-step optimization problem:

$$\begin{aligned}
&\max_{\kappa_m,\kappa_{e,i},\kappa_{f,i,j}} \mathrm{E}\{\max_{\mathbf{X}_1} \mathrm{E}[W_2 \mid \mathrm{F}_1] - \frac{\gamma}{2}\mathrm{V}[W_2 \mid \mathrm{F}_1]\},\\
&s.t. \quad W_2 = W_1 + \mathbf{X}_1^T(\mathbf{V} - \mathbf{P}_1),\\
&\qquad \kappa_m + \sum_{i=1}^{n}\kappa_{e,i} + \sum_{i=1}^{n}\sum_{j=1}^{k}\kappa_{f,i,j} \le K\\
&\qquad \kappa_m \ge 0, \kappa_{e,i} \ge 0, \kappa_{f,i,j} \ge 0, \forall i, j,
\end{aligned} \tag{10}$$

where F_1 contains the information in s_m, $s_{e,i}$ and $s_{f,i,j}$ for all i and j. The investor firstly chooses $\kappa_m, \kappa_{e,i}$ and $\kappa_{f,i,j}$ to allocate her attention optimally, and maximizes the expected utility by choosing the optimal quantity $\mathbf{X}_1$ conditioning on F_1.

The optimization problem above can be solved using backward induction. At time 1 given the attention allocation κ s, the optimal demand $\mathbf{X}_1$ is available in closed form since the objective function is just a quadratic function of $\mathbf{X}_1$:

$$\mathbf{X}_1 = \frac{1}{\gamma}\mathrm{V}[\mathbf{V} \mid \mathrm{F}_1]^{-1}(\mathrm{E}[\mathbf{V} \mid \mathrm{F}_1] - \mathbf{P}_1).$$

In the equilibrium, the market clearing condition is $\mathbf{X}_1 = 1_{nk\times 1}$, where 1_{nk} is a nk -by-1 vector of 1s. Substituting this into the equation above yields:

$$\mathbf{P}_1 = \mathrm{E}[\mathbf{V} \mid \mathrm{F}_1] - \gamma \mathrm{V}[\mathbf{V} \mid \mathrm{F}_1] 1_{nk\times 1} \tag{11}$$

Substituting $\mathbf{X}_1 = 1_{nk\times 1}$ and the above Equation into Equation (10), the optimization problem reduces to:

$$\begin{aligned}&\max_{\kappa_m,\kappa_{e,i},\kappa_{f,i,j}} -\frac{\gamma}{2}\mathrm{V}[\mathbf{V}\,|\,\mathrm{F}_1],\\ &s.t.\quad \kappa_m+\sum_{i=1}^{n}\kappa_{e,i}+\sum_{i=1}^{n}\sum_{j=1}^{k}\kappa_{f,i,j}\le K,\\ &\qquad \kappa_m\ge 0,\kappa_{e,i}\ge 0,\kappa_{f,i,j}\ge 0,\forall i,j,n.\end{aligned} \tag{12}$$

Note that $\mathrm{V}[m\,|\,\mathrm{F}_1]=\mathrm{V}[m\,|\,s_f]=1/(\tau+\kappa_m)$, which is simply the variance of the posterior distribution of m taking the unconditional distribution of m as a prior. Similarly, $\mathrm{V}[e_i\,|\,\mathrm{F}_1]=1/(\tau+\kappa_{e,i})$ and $\mathrm{V}[f_{i,j}\,|\,\mathrm{F}_1]=1/(\tau+\kappa_{f,i,j}E_{i,j})$. The optimization problem in Equation (12) is equivalent to the following:

$$\begin{aligned}&\min_{\kappa_m,\kappa_{e,i},\kappa_{f,i,j}} \frac{n^2k^2}{\tau+\kappa_m}+\sum_{i=1}^{n}\frac{k^2}{\tau+\kappa_{e,i}}+\sum_{i=1}^{n}\sum_{j=1}^{k}\frac{1}{\tau+\kappa_{f,i,j}E_{i,j}}\\ &s.t.\quad \kappa_m+\sum_{i=1}^{n}\kappa_{e,i}+\sum_{i=1}^{n}\sum_{j=1}^{k}\kappa_{f,i,j}\le K\\ &\qquad \kappa_m\ge 0,\kappa_{e,i}\ge 0,\kappa_{f,i,j}\ge 0,\forall i,j,\end{aligned}$$

This optimization problem can be solved via the Lagrangian multiplier:

$$\begin{aligned}L=&\frac{n^2k^2}{\tau+\kappa_m}+\sum_{i=1}^{n}\frac{k^2}{\tau+\kappa_{e,i}}+\sum_{i=1}^{n}\sum_{j=1}^{k}\frac{1}{\tau+\kappa_{f,i,j}E_{i,j}}\\ &-\lambda_K\Big(K-\kappa_m-\sum_{i=1}^{n}\kappa_{e,i}-\sum_{i=1}^{n}\sum_{j=1}^{k}\kappa_{f,i,j}\Big)\\ &-\lambda_m\kappa_m-\sum_{i=1}^{n}\lambda_{e,i}\kappa_{e,i}-\sum_{i=1}^{n}\sum_{j=1}^{k}\lambda_{f,i,j}\kappa_{f,i,j}\end{aligned}$$

The above function is a strictly decreasing and convex function of κ s, therefore it must have one unique solution. The solution can be easily derived by setting the first order conditions of the above Langrangian multiplier to zero and solve for the system of inequalities. After some calculation, we obtain:

$$\kappa_m = \frac{nk}{\sqrt{\lambda_K}} - \tau,$$
$$\kappa_{e,i} = \frac{k}{\sqrt{\lambda_K}} - \tau,$$
$$\kappa_{f,i,j} = \frac{1}{\sqrt{E_{i,j}\lambda_K}} - \frac{\tau}{E_{i,j}},$$
$$\frac{1}{\sqrt{\lambda_K}} = \frac{K + (1+n+C_1)\tau}{2nk + C_2},$$
$$\lambda_m = \lambda_{e,i} = \lambda_{f,i,j} = 0,$$

in which $C_1 = \sum_{i=1}^{n}\sum_{j=1}^{k} 1/E_{i,j}$ and $C_2 = \sum_{i=1}^{n}\sum_{j=1}^{k} 1/\sqrt{E_{i,j}}$ are constants that can be interpreted as overall measures of information processing cost for all stocks. Since we are not interested in the scenario when there is a scarcity of the attention resources, we assume that K is sufficiently large, so that we always have $\kappa_m > \kappa_e > 0$, which is a result documented in Huang *et al*., (2018).

1.4 Social media coverage, financial reporting opacity and return co-movement

We proceed to demonstrate the relationship between social media coverage, financial reporting opacity and return co-movements. We firstly define the firm specific, industry and market returns from time 0 to time 1 as follows:

$$r_{i,j} = P_{i,j,1} - P_{i,j,0},$$
$$r_i = \frac{1}{k}\sum_{j=k}^{n} r_{i,j},$$
$$r_m = \frac{1}{nk}\sum_{i=1}^{n}\sum_{j=1}^{k} r_{i,j}.$$

So the industry returns and market returns are just equally weighted returns for the stocks. Also, our results do not differ qualitatively when using other weighting schemes with fixed weights.

The price of each stock at time 1 is given by Equation (11). Note that the conditional variance part $\mathrm{V}[\mathbf{V} \mid \mathrm{F}_1]$ is non-random. We can therefore write the return $r_{i,j}$ as:

$$r_{i,j} = C + \frac{\kappa_m}{\tau + \kappa_m} s_m + \frac{\kappa_e}{\tau + \kappa_e} s_{e,i} + \frac{E_{i,j}\kappa_{e,i}}{\tau + E_{i,j}\kappa_{e,i}} s_{f,i,j},$$

where $C = -\gamma(\frac{nk}{\tau + \kappa_m} + \frac{nk}{\tau + \kappa_e} + \sum_{i=1}^{n}\sum_{j=1}^{k}\frac{1}{\tau + E_{i,j}\kappa_{f,i,j}}) - P_0$ is a constant. Our primary measures of the co-movement between market return, industry return and firm-specific return are the R^2 obtained from regressing firm-specific return on market and industry returns correspondingly. This can be interpreted as the variation in $r_{i,j}$ that can be explained by r_m and r_i respectively. We can decompose the variance of $r_{i,j}$ into three components: variance due to market shock Var_m, industry shock Var_i and firm-specific shock $Var_{i,j}$. These three components are formally defined as follows:

$$Var_m = \mathrm{V}\left[\frac{\kappa_m}{\tau + \kappa_m} s_m\right] = 1 - \frac{\tau\sqrt{\lambda_K}}{nk},$$
$$Var_i = \mathrm{V}\left[\frac{\kappa_e}{\tau + \kappa_e} s_{e,i}\right] = 1 - \frac{\tau\sqrt{\lambda_K}}{k},$$
$$Var_{i,j} = \mathrm{V}\left[\frac{E_{i,j}\kappa_{e,i}}{\tau + E_{i,j}\kappa_{e,i}} s_{f,i,j}\right] = 1 - \frac{\tau\sqrt{\lambda_K}}{\sqrt{E_{i,j}}}.$$

Notice that we are only interested in the stocks with $\sqrt{\frac{\lambda_K}{E_{i,j}}} < \frac{1}{\tau}$, since when this does not hold, $\kappa_{f,i,j} = 0$ and the co-movement is irrelevant of the information processing cost. The market and industry R^2 measures are then formally defined as follows:

$$R^2_{i,j,m} = \frac{Var_m}{Var_m + Var_i + Var_{i,j}},$$
$$R^2_{i,j,i} = \frac{Var_m + Var_i}{Var_m + Var_i + Var_{i,j}}.$$

Differentiating w.r.t. $E_{i,j}$ yields:

$$\frac{\partial R^2_{i,j,m}}{\partial E_{i,j}} = -\frac{Var_m \tau \sqrt{\lambda_K}}{2(Var_m + Var_i + 1 - \frac{\tau\sqrt{\lambda_K}}{\sqrt{E_{i,j}}})^2 E^{3/2}_{i,j}},$$
$$\frac{\partial R^2_{i,j,i}}{\partial E_{i,j}} = -\frac{(Var_m + Var_i)\tau \sqrt{\lambda_K}}{2(Var_m + Var_i + 1 - \frac{\tau\sqrt{\lambda_K}}{\sqrt{E_{i,j}}})^2 E^{3/2}_{i,j}}. \tag{13}$$

We deduce Proposition 3 and 4 from the above relationships, which leads directly to Propositions 1 and 2 stated in Section 2 respectively:

Proposition 3: The co-movement measures $R^2_{i,j,m}$ and $R^2_{i,j,i}$ decrease monotonically as $\alpha_{i,j}$ increases, holding $O_{i,j}$ constant.

Proof. *Since* $E_{i,j} = \alpha_{i,j} + O_{i,j}$, *it is obvious from (13) that the derivatives of both co-movement measures w.r.t. social media coverage are strictly negative.* ■

Proposition 4: The marginal effects of an increase in social media coverage on the co-movement measures $R^2_{i,j,m}$ and $R^2_{i,j,i}$ are larger for firms with smaller $O_{i,j}$, or higher financial reporting opacity.

Proof. *We only need to show that the denominators in* $\frac{\partial R^2_{i,j,m}}{\partial E_{i,j}}$ *and* $\frac{\partial R^2_{i,j,i}}{\partial E_{i,j}}$ *increases as a function of* $O_{i,j}$, *which evidently holds.* ■

It is evident that Propositions 3 and 4 lead to Propositions 1 and 2 directly. Moreover, we can test for these propositions empirically by using fixed-effect panel regressions, since the coefficients of the explanatory variables and interaction terms can be interpreted naturally as marginal effects.

Appendix 2: Seeking Alpha article example

Amazon Earnings Broadly As Expected

Apr. 25, 2013 5:39 PM ET
135 comments About: Amazon.com, Inc. (AMZN)
Paulo Santos
(10,045 followers)
Long/short equity, arbitrage, event-driven

Amazon (NASDAQ:AMZN) reported its Q1 2013 earnings. These came in at $0.18 versus a $0.09 consensus. At first the stock climbed quite a bit on the notion that it had beat or doubled expectations, but one needs to consider that for Amazon, $0.10 in excess or missing on its earnings is basically irrelevant, because it needs just $46 million or a puny 0.28% of sales for a beat or miss of that magnitude.

Also predictable, Amazon's revenues came in slightly below consensus ($16.07 billion vs $16.16 billion consensus). More relevant was Amazon's guidance for Q2 2013, which was as follows:

Net sales are expected to be between $14.5 billion and $16.2 billion, or to grow between 13% and 26% compared with second quarter 2012.

Operating income (loss) is expected to be between $(340) million and $10 million, compared to $107 million in the comparable prior year period.

This guidance includes approximately $340 million for stock-based compensation and amortization of intangible assets, and it assumes, among other things, that no additional business acquisitions, investments, or legal settlements are concluded and that there are no further revisions to stock-based compensation estimates.

As I predicted before the earnings were released, this constitutes another guide-down for Amazon's revenues. The midpoint of the guidance falls at $15.35 billion whereas present consensus sits at $15.94. I'd expect consensus to be revised lower to around $15.7-$15.8 billion or so.

Comparison to my model 1

The model 1 predictions compared as follows to what Amazon actually reported:

($ million)	Model 1 Q1 2013	Actual Q1 2013
Revenues		
Product	13161	13271
Services	2782	2799
total revenues	**15943**	**16070**
COGS product	11714	11801
Gross margin	4229	4269
Gross margin % of revenues	**26.5%**	**26.6%**
Fulfillment	1716	1796
Marketing	624	632
Technology	1464	1383
G&A	265	246
Other	25	32
total operating costs	**15807**	**15890**
Operating income	**136**	**180**
% of revenues	**0.85%**	**1.12%**
Interest income	10	10
Interest expense	-28	-33
Other	-20	-77
total	**-38**	**-100**
Income before taxes	**98**	**80**
Income taxes	-34	18
Equity-method	-25	-17
Net income	**38**	**82**
Diluted shares (million)	463.0	463.0
EPS	**$0.08**	**$0.18**

* Using gross margin from GMV

In what regards my own modelling, where I use my model 1 for both short term and long term predictions, the major differences were in 3 cost lines and 1 margin line:

- **Product margins** came in at 11.3% versus my 11.0% assumption. My 2013 assumption is 11.1%, which I will revise towards 11.2%;
- **Technology**, which came in 5.5% below my estimate. This implied a ratio of Technology/Other revenue of 173.3% versus my Q1 2013 assumption of 183% … but it should be

noted that my 2013 assumption is 175% so lower than Q1 2013. I will revise my long term assumptions down 2% per year as a result;

- **G&A**, which came in 7.2% below my estimate. This implied a ratio of G&A/GMV of 0.78% versus my assumption of 0.85%. Since my 2013 assumption is already 0.80% this will mean no change as this number is somewhat volatile and the yearly assumption is already below Q1 and near the realised value;
- **Fulfillment**, which came in 4.7% above my estimate. This implied a ratio of Fulfillment/GMV of 5.71% versus my assumption of 5.50%. Q1 is usually the lowest in this regard so this implies a higher 2013 assumption. Presently the assumption is at 5.64%, so I will change the model towards 5.7%.

My own long-term model already implies that technology will get better (less costly) over the long-term, so no surprise there. G&A has some volatility so it won't imply much of a change. As for fulfilment, it might have negative implications for the long term.

All in all the cost relationships held quite well. The minor $40 million difference in net profit is well within the kind of uncertainty one can expect while predicting a company of Amazon's size and basically came from the product margins being slightly ahead of expectations, probably still from the higher margins enjoyed by the new Kindle Fires.

It should also be noted that every revenue growth assumption was very close to what Amazon reported, from 1P to 3P to other revenue.

Revised long-term model

Taking into account the slight differences explained, my revised long-term model now predicts the following:

Amazon.com P&L ($ million)

	2012	2013	2014	2015	2016	2017	2018	2019	2020
Revenues		20.4%	18.7%	16.7%	14.4%	13.1%	11.6%	10.0%	8.8%
Product	51733	60528	69607	78656	87308	96039	104682	113057	120971
Services	9360	13004	17687	23224	29291	35868	42592	48980	55348
total revenues	61093	73531	87293	101880	116598	131907	147274	162037	176318
COGS product	45971	53749	61811	69846	77529	85282	92958	100394	107422
Gross margin using GMV	15124	19375	24068	29353	35084	41213	47500	53636	59693
Gross margin	15122	19783	25483	32034	39069	46625	54316	61643	68896
Gross margin % of revenues	24.8%	26.9%	29.2%	31.4%	33.5%	35.3%	36.9%	38.0%	39.1%
Fulfillment	6419	8181	10162	12394	14813	17401	20056	22646	25204
Marketing	2408	3014	3744	4566	5458	6411	7389	8343	9286
Technology	4564	6763	9558	12560	15794	19076	22091	24484	26627
G&A	896	1148	1426	1739	2079	2442	2815	3178	3537
Other	159	100	100	100	100	100	100	100	100
total operating costs	60417	72954	86800	101205	115773	130712	145408	159146	172175
Operating income	676	577	493	675	825	1195	1866	2891	4143
% of revenues	1.11%	0.79%	0.56%	0.66%	0.71%	0.91%	1.27%	1.78%	2.35%
Interest income	40								
Interest expense	-92								
Other	-80	-150							
total	-132	-200	-200	-200	-200	-200	-200	-200	-200
Income before taxes	544	377	293	475	625	995	1666	2691	3943
Income taxes	-428	-82	-102	-166	-219	-348	-583	-942	-1380
Equity-method	-155	-100	-100	-100	-100	-100	-100	-100	-100
Net income	-39	195	90	209	306	547	983	1649	2463
Diluted shares (million)	470.0	473.8	478.4	483.6	489.7	496.5	504.2	512.6	521.8
EPS	-$0.08	$0.41	$0.19	$0.43	$0.63	$1.10	$1.95	$3.22	$4.72

-++

The predictions are unchanged for the most part, with the margins and cost lines basically compensating each other, only the lower tax rate ends up having a slight positive effect for 2013.

Conclusion

Amazon's earnings report brought nothing new. The growth rates and costs continue mostly as expected - the cost relationships held, with most uncertainty remaining on technology, where improvement is already expected and always difficult to model.

These cost relationships mean that Amazon will have a lot of trouble ever meeting the lofty expectations the Street has for it. At the same time Amazon's growth rates continue to falter and perhaps somewhat amazingly, net shipping costs increased again.

There was nothing in the report to change my opinion that Amazon is a clear short which will never produce enough profit to justify the levels it trades at. I expect this report to lead to another round of downward estimate revisions in terms of revenues, and perhaps also in terms of EPS. These revisions are systematic because the long-term models the Street uses do not respect the stable cost relationships that I have identified,

Disclosure: I am short AMZN. I wrote this article myself, and it expresses my own opinions. I am not receiving compensation for it (other than from Seeking Alpha). I have no business relationship with any company whose stock is mentioned in this article.

Appendix 3: Variable definitions

Variable	Definition
SYNCH	Stock return synchronicity after log transformation
CORRE	Percentage points of time-series Pearson correlation coefficient between weekly firm return and weekly market returns
L_SA	Natural log of (SA article number + 1)
LNUM	Natural log of (number of analyst coverage + 1)
SIZE	Natural log of (Firm's market capitalisation)
LMB	Natural log of (market capitalisation scaled by the book value of equity)
LEVERAGE	Total long term debt scaled by total assets
ROA	Income before extraordinary items scaled by total assets
NIND	Natural log of the number of firms in the industry to which firm I belongs
HERFSALE	Sum of squared terms of the proportion of a firm's revenue to total revenue in the industry
STDROA	Standard deviation of the ratio between income before extraordinary items and total asset in the previous five years
BIG4	Dummy variable, set to 1 if the firm is audited by one of the Big 4 audit firms
OPA	Measure of financial opacity, which is the average of accrual quality value of the previous three years. Accrual quality measure is based on the Francis *et al.*, (2005) model. Measured as absolute value of residual from each year two-digit SIC industry cross sectional regression: $\frac{TAcc_{i,t}}{TA_{i,t-1}} = \alpha_0 + \frac{\beta_1 CFO_{i,t-1}}{TA_{i,t-1}} + \frac{\beta_2 CFO_{i,t}}{TA_{i,t-1}} + \frac{\beta_3 CFO_{i,t+1}}{TA_{i,t-1}} + \frac{\beta_4\, \Delta Rev_{i,t}}{TA_{i,t-1}} + \frac{\beta_5\, PPE_{i,t}}{TA_{i,t-1}} + \varepsilon_{i,t}$
OPA2	Analyst earnings forecast dispersion scaled by firm's opening stock price of the year
EARNING	Earnings before extraordinary items scaled by total assets
L_SA_IND	Natural log of (the total annual SA coverage of the industry to which the firm belongs + 1)
L_ADX	Natural log of (firm's annual advertising expenditure + 1)

Figure 1: Summary statistics of Seeking Alpha articles

Year	**Seeking Alpha article**	**Firms covered by Seeking Alpha**
2004	27	19
2005	1165	426
2006	5753	1707
2007	13710	2809
2008	13111	2721
2009	15842	2645
2010	13996	2808
2011	13719	2813
2012	14188	2818
2013	16767	3089
2014	24939	3969

Table 1: Descriptive statistics

Variable	Mean	Std. Dev.	Min	P25	Median	P75	Max	N
SYNCH	-2.664	1.962	-16.689	-3.874	-2.421	-1.234	5.508	39568
CORRE	40.898	23.194	-45.684	24.361	43.003	58.662	94.502	39568
L_SA	0.426	0.788	0.000	0.000	0.000	0.693	6.632	39568
LNUM	0.952	1.108	0.000	0.000	0.000	1.923	4.022	39568
SIZE	6.213	2.119	-1.038	4.661	6.184	7.670	13.348	39568
LMB	0.702	0.883	-3.332	0.176	0.634	1.150	9.241	39568
LEVERAGE	0.154	0.173	0.000	0.001	0.097	0.255	0.969	39568
ROA	0.025	2.308	-10.655	-0.008	0.024	0.067	226.310	39568
NIND	5.157	1.195	0.000	4.263	5.389	6.194	6.796	39568
HERFSALE	0.075	0.080	0.010	0.035	0.045	0.090	1.000	39568
STDROA	0.101	0.960	0.000	0.013	0.033	0.087	92.564	39568
BIG4	0.701	0.458	0.000	0.000	1.000	1.000	1.000	39568
OPA	0.071	0.156	0.000	0.024	0.042	0.080	8.641	22847

Table 1 presents the descriptive statistics of variables used in the analysis. The sample contains 39,568 firm-year observations over the period 2004–2014. P25 (P75) is the 25th (75th) percentile of the variable's distribution. SYNCH, stock price synchronicity, is defined as the log-transformation of the adjusted R^2 of the firm-year estimation regressing weekly stock return on weekly market- and industry-level returns; CORRE is the percentage points of time-series Pearson correlation coefficient between weekly firm returns and the weekly market returns; L_SA is the natural log of 1 plus the number of single-ticker Seeking Alpha articles for a firm in the year; LNUM is the natural log of 1 plus the number of analysts following the firm for the previous fiscal year; SIZE is the natural log of the firm's market capitalisation at the end of the last fiscal year; LMB is the natural log of market capitalisation scaled by the book value of equity at the end of the last fiscal year; LEVERAGE is the total long-term debt scaled by total assets at the end of the last fiscal year; ROA is income before extraordinary items divided by total assets at the end of the last fiscal year; NIND is the natural log of the number of firms in the industry to which firm *i* belongs at the end of the last fiscal year; HERFSALE is the sum of the squared terms of the proportion of a firm's revenue to total revenue in the industry at the end of the last fiscal year; STDROA is the standard deviation of the ratio between income before extraordinary items and total assets in the previous five years; BIG4 is a dummy variable that equals 1 if the firm is audited by one of the Big 4 audit firms, zero otherwise; OPA is the previous five years' standard deviation of accrual quality value based on the Francis *et al.* (2005) model.

Table 2: Correlation table

	SYNCH	CORRE	L_SA	LNUM	SIZE	LMB	LEVERAGE	ROA	NIND	HERFSALE	STDROA
CORRE	0.847***										
L_SA	0.262***	0.222***									
LNUM	0.41***	0.393***	0.398***								
SIZE	0.63***	0.592***	0.464***	0.541***							
LMB	0.085***	0.086***	0.143***	0.154***	0.336***						
LEV	0.195***	0.171***	0.058***	0.099***	0.235***	0.083***					
ROA	0.018***	0.019***	0.015***	0.006	0.037***	0.115***	-0.010				
NIND	-0.153***	-0.109***	-0.1***	-0.134***	-0.127***	0.028***	-0.181***	0.011*			
HERFSALE	0.090***	0.061***	0.065***	0.095***	0.053***	-0.028***	0.082***	-0.006	-0.74***		
STDROA	-0.031***	-0.027***	0.003	-0.035***	-0.035***	0.131***	-0.036***	0.814***	0.03***	-0.017***	
BIG4	0.408***	0.402***	0.203***	0.315***	0.544***	0.161***	0.197***	0.019***	-0.167***	0.086***	-0.005

Table 2 reports the Pearson correlation of variables used in the analysis. The sample contains 39,568 firm-year observations over the period 2004–2014. SYNCH, stock price synchronicity, is defined as the log-transformation of the adjusted R^2 of the firm-year estimation regressing weekly stock return on weekly market- and industry-level returns; CORRE is the percentage points of time-series Pearson correlation coefficient between weekly firm return and weekly market returns; L_SA is the natural log of 1 plus the number of single-ticker Seeking Alpha articles for a firm in the year; LNUM is the natural log of 1 plus the number of analysts following the firm for the previous fiscal year; SIZE is the natural log of the firm's market capitalisation at the end of the last fiscal year; LMB is the natural log of market capitalisation scaled by the book value of equity at the end of the last fiscal year; LEVERAGE is the total long-term debt scaled by total assets at the end of the last fiscal year; ROA is income before extraordinary items divided by total assets at the end of the last fiscal year; NIND is the natural log of the number of firms in the industry to which firm *i* belongs at the end of the last fiscal year; HERFSALE is the sum of the squared terms of the proportion of a firm's revenue to total revenue in the industry at the end of the last fiscal year; STDROA is the standard deviation of the ratio between income before extraordinary items and total assets in the previous five years; BIG4 is a dummy variable that equals 1 if the firm is audited by one of the Big 4 audit firms, zero otherwise; *, ** and *** denote significance levels at 10%, 5% and 1% respectively.

Table 3: Baseline Model

$$Comovement_{i,t} = \alpha_0 + \beta_1 L\_SA_{i,t} + \beta_2 LNUM_{i,t} + \beta_3 SIZE_{i,t} + \beta_4 LMB_{i,t} + \beta_5 LEVERAGE_{i,t} + \beta_6 ROA_{i,t}$$
$$+\beta_7 NIND_{i,t} + \beta_8 HERFSALE_{i,t} + \beta_9 STDROA_{i,t} + \beta_{10} BIG4_{i,t} + \sum \alpha_i Firm_i + \sum \alpha_j Year_j + \varepsilon_{i,t}$$

VARIABLES	SYNCH	SYNCH	CORRE	CORRE
L_SA	-0.055***	-0.059***	-0.450**	-0.493***
	[-3.852]	[-4.137]	[-2.364]	[-2.585]
LNUM	0.141***	0.122***	1.153***	0.972***
	[7.128]	[6.083]	[4.448]	[3.683]
SIZE	0.373***	0.444***	5.332***	5.909***
	[23.917]	[22.109]	[27.081]	[23.157]
LMB		-0.142***		-1.158***
		[-6.701]		[-4.447]
LEV		0.160		2.444**
		[1.624]		[2.011]
ROA		-0.037		-0.148
		[-0.525]		[-0.174]
NIND		-0.036		0.671
		[-0.353]		[0.535]
HERFSALE		0.411		-7.096
		[0.807]		[-1.078]
STDROA		0.015		0.574
		[0.133]		[0.440]
BIG4		0.061		0.603
		[1.471]		[1.189]
CONSTANT	-5.273***	-5.474***	3.369***	-2.920
	[-53.159]	[-9.843]	[2.706]	[-0.424]
Observations	39,568	39,568	39,568	39,568
R-squared	0.704	0.705	0.698	0.698
firm fixed effects	Yes	Yes	Yes	Yes
time fixed effects	Yes	Yes	Yes	Yes
Adj. R-squared	0.648	0.649	0.641	0.641

Table 3 presents the effect of Seeking Alpha coverage on stock return co-movement. The sample contains 39,568 firm-year observations over the period 2004–2014. SYNCH, stock price synchronicity, is defined as the log-transformation of the adjusted R^2 of the firm-year estimation regressing weekly stock return on weekly market- and industry-level returns; CORRE is the percentage points of time-series Pearson correlation coefficient between weekly firm return and weekly market returns L_SA is the natural log of 1 plus the number of single-ticker Seeking Alpha articles for a firm in the year; LNUM is the natural log of 1 plus the number of analysts following the firm for the previous fiscal year; SIZE is the natural log of the firm's market capitalisation at the end of the last fiscal year; LMB is the natural log of market capitalisation scaled by the book value of equity at the end of the last fiscal year; LEVERAGE is the total long-term debt scaled by total assets at the end of the last fiscal year; ROA is income before extraordinary items divided by total assets at the end of the last fiscal year; NIND is the natural log of the number of firms in the industry to which firm i belongs at the end of the last fiscal year; HERFSALE is the sum of the squared terms of the proportion of a firm's revenue to total revenue in the industry at the end of the last fiscal year; STDROA is the standard deviation of the ratio between income before extraordinary items and total assets in the previous five years; BIG4 is a dummy variable that equals 1 if the firm is audited by one of the Big 4 audit firms, zero otherwise; T-statistics robust to heteroscedasticity and clustered by firm are reported in the square brackets; *, ** and *** denote significance levels at 10%, 5% and 1% respectively.

Table 4: Results relating to moderating effect of financial reporting opacity

$$Comovement_{i,t} = \alpha_0 + \beta_1 L\_SA_{i,t} + \beta_2 OPA_{i,t} + \beta_3 L\_SA_{i,t} * OPA_{i,t} + \beta_4 LNUM_{i,t} + \beta_5 SIZE_{i,t} + \beta_6 LMB_{i,t} + \beta_7 LEVERAGE_{i,t}$$
$$+\beta_8 ROA_{i,t} + \beta_9 NIND_{i,t} + \beta_{10} HERFSALE_{i,t} + \beta_{11} STDROA_{i,t} + \beta_{12} BIG4_{i,t} + \sum \alpha_i Firm_i + \sum \alpha_j Year_j + \varepsilon_{i,t}$$

VARIABLES	SYNCH	CORRE	SYNCH	CORRE
L_SA	-0.050***	-0.210	-0.070***	-0.541*
	[-2.807]	[-0.883]	[-3.379]	[-1.836]
OPA	-0.012	-0.460		
	[-0.109]	[-0.392]		
OPA2			0.000***	0.000
			[3.616]	[0.987]
L_SA*OPA	-0.143***	-1.939***		
	[-4.294]	[-3.637]		
L_SA*OPA2			-0.071***	-0.435**
			[-4.561]	[-2.453]
LNUM	0.117***	0.993***	0.071**	0.322
	[4.413]	[2.957]	[2.195]	[0.706]
SIZE	0.412***	5.203***	0.359***	3.594***
	[15.101]	[15.196]	[9.788]	[7.419]
LMB	-0.138***	-0.806**	-0.140***	-0.666
	[-4.898]	[-2.366]	[-3.763]	[-1.311]
LEV	0.163	3.147**	0.230	2.941
	[1.276]	[1.973]	[1.498]	[1.438]
ROA	0.030	1.394	0.228	3.500**
	[0.331]	[1.322]	[1.639]	[2.052]
NIND	-0.324**	0.015	-0.303*	1.591
	[-2.166]	[0.008]	[-1.908]	[0.670]
HERFSALE	0.989	8.704	-0.366	-12.894
	[1.288]	[0.876]	[-0.453]	[-1.135]
STDROA	-0.009	2.033	-0.119	-1.905
	[-0.057]	[1.160]	[-0.493]	[-0.691]
BIG4	0.069	0.942	0.083	0.955
	[1.151]	[1.254]	[1.083]	[0.974]
CONSTANT	-3.810***	3.226	-3.212***	12.313
	[-4.668]	[0.296]	[-3.719]	[0.976]
Observations	22,847	22,847	12,606	12,606
firm fixed effects	Yes	Yes	Yes	Yes
time fixed effects	Yes	Yes	Yes	Yes
Adj. R-squared	0.642	0.638	0.617	0.553

Table 4 presents the results of varying effect of Seeking Alpha coverage on stock return co-movement with respect to Opacity. The sample contains 22,847 (12,606) firm-year observations over the period 2004–2014 when SYNCH and CORRE are the dependent variable.[17] SYNCH, stock price synchronicity, is defined as the log-transformation of the adjusted R^2 of the firm-year estimation regressing weekly stock return on weekly market- and industry-level returns; CORRE is the percentage points of time-series Pearson correlation coefficient between weekly firm return and weekly market returns; L_SA is the natural log of 1 plus the number of single-ticker Seeking Alpha articles for a firm in the year; OPA is the previous five years standard deviation of Accrual quality value based on Francis et al (2005) model; OPA2 is the analyst earnings forecast dispersion scaled by the firm's opening stock price of the fiscal year; LNUM is the natural log of 1 plus the number of analysts following the firm for the previous fiscal year; SIZE is the natural log of the firm's market capitalisation at the end of the last fiscal year; LMB is the natural log of market capitalisation scaled by the book value of equity at the end of the last fiscal year; LEVERAGE is the total long-term debt scaled by total assets at the end of the last fiscal year; ROA is income before extraordinary items divided by total assets at the end of the last fiscal year; NIND is the natural log of the number of firms in the industry to which firm i belongs at the end of the last fiscal year; HERFSALE is the sum of the squared terms of the proportion of a firm's revenue to total revenue in the industry at the end of the last fiscal year; STDROA is the standard deviation of the ratio between income before extraordinary items and total assets in the previous five years; BIG4 is a dummy variable that equals 1 if the firm is audited by one of the Big 4 audit firms, zero otherwise. T-statistics robust to heteroscedasticity and clustered by firm are reported in the square brackets. *, ** and *** denote significance levels at 10%, 5% and 1% respectively.

[17] The sample size is reduced due to the availability of data required to calculate the variable OPA and OPA2.

Table 5: Propensity score matching

VARIABLES	SYNCH	CORRE
L_SA	-0.055***	-1.133***
	[-3.138]	[-5.111]
LNUM	0.083***	0.772**
	[2.956]	[2.208]
SIZE	0.370***	4.309***
	[13.106]	[12.425]
LMB	-0.116***	-0.874***
	[-4.142]	[-2.629]
LEVERAGE	0.165	2.906
	[1.144]	[1.631]
ROA	0.004**	0.082***
	[2.211]	[3.782]
NIND	-0.045	-3.240*
	[-0.329]	[-1.691]
HERFSALE	-1.461**	-28.209***
	[-2.302]	[-3.681]
STDROA	-0.069***	-0.424***
	[-7.613]	[-3.853]
BIG4	0.010	-0.602
	[0.134]	[-0.644]
CONSTANT	-4.863***	29.478***
	[-6.120]	[2.775]
Observations	21,528	21,528
Firm fixed effects	Yes	Yes
Year fixed effects	Yes	Yes
Adj. R-squared	0.661	0.629

Table 5 presents the effect of Seeking Alpha coverage on stock return co-movement for PSM matched sample. The sample contains 21,528 firm-year observations over the period 2004–2014. SYNCH, stock price synchronicity, is defined as the log-transformation of the adjusted R^2 of the firm-year estimation regressing weekly stock return on weekly market- and industry-level returns; CORRE is the percentage points of time-series Pearson correlation coefficient between weekly firm return and weekly market returns; L_SA is the natural log of 1 plus the number of single-ticker Seeking Alpha articles for a firm in the year; LNUM is the natural log of 1 plus the number of analysts following the firm for the previous fiscal year; SIZE is the natural log of the firm's market capitalisation at the end of the last fiscal year; LMB is the natural log of market capitalisation scaled by the book value of equity at the end of the last fiscal year; LEVERAGE is the total long-term debt scaled by total assets at the end of the last fiscal year; ROA is income before extraordinary items divided by total assets at the end of the last fiscal year; NIND is the natural log of the number of firms in the industry to which firm i belongs at the end of the last fiscal year; HERFSALE is the sum of the squared terms of the proportion of a firm's revenue to total revenue in the industry at the end of the last fiscal year; STDROA is the standard deviation of the ratio between income before extraordinary items and total assets in the previous five years; BIG4 is a dummy variable that equals one if the firm is audited by one of the Big 4 audit firms, zero otherwise. Both firm and year fixed effects are used. T-statistics robust to heteroscedasticity and clustered by firm are reported in the square brackets. *, ** and *** denote significance levels at 10%, 5% and 1% respectively.

Table 6: Two-stage-least-square (2SLS)

VARIABLES	First stage	Second stage SYNCH	Second stage CORRE
L_SA_IND	0.029***		
	[11.617]		
L_ADX	0.038***		
	[3.058]		
L_SA_P		-0.472**	-11.544***
		[-2.450]	[-4.478]
LNUM	0.043***	0.141***	1.480***
	[3.534]	[6.688]	[4.848]
SIZE	0.119***	0.496***	7.284***
	[12.018]	[16.306]	[17.903]
LMB	-0.035***	-0.159***	-1.613***
	[-3.345]	[-7.478]	[-5.696]
LEVERAGE	0.216***	0.251**	4.886***
	[4.948]	[2.480]	[3.605]
ROA	0.019	-0.032	0.004
	[0.709]	[-0.481]	[0.005]
NIND	0.166***	0.040	2.705**
	[3.377]	[0.390]	[1.999]
HERFSALE	0.773**	0.715	1.023
	[2.532]	[1.443]	[0.143]
STDROA	0.127***	0.071	2.054
	[2.621]	[0.649]	[1.571]
BIG4	0.056***	0.084**	1.227**
	[3.959]	[2.119]	[2.435]
Observations	38,821	38,821	38,821
Firm fixed effects	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes
Adj. R-squared	0.592	0.640	0.588

Table 6 presents the two-stage least square (2SLS) analysis results. We use two instrumental variables: L_ADX is the natural log of 1 plus the firm's annual advertisement expenditure; L_SA_IND is the natural log of 1 plus the total annual SA coverage of the industry to which the firm belongs. The sample contains 38,821 firm-year observations over the period 2004–2014. SYNCH, stock price synchronicity, is defined as the log-transformation of the adjusted R^2 of the firm-year estimation regressing weekly stock return on weekly market- and industry level return; CORRE is the percentage points of time-series Pearson correlation coefficient between weekly firm return and weekly market returns; L_SA is the natural log of 1 plus the number of single-ticker Seeking Alpha articles for a firm in the year; L_SA_P is the predicted value of L_SA from the first stage. LNUM is the natural log of 1 plus the number of analysts following the firm for the previous fiscal year; SIZE is the natural log of the firm's market capitalisation at the end of the last fiscal year; LMB is the natural log of market capitalisation scaled by the book value of equity at the end of the last fiscal year; LEVERAGE is the total long-term debt scaled by total assets at the end of the last fiscal year; ROA is income before extraordinary items divided by total assets at the end of the last fiscal year; NIND is the natural log of the number of firms in the industry to which firm i belongs at the end of the last fiscal year; HERFSALE is the sum of the squared terms of the proportion of a firm's revenue to total revenue in the industry at the end of the last fiscal year; STDROA is the standard deviation of the ratio between income before extraordinary items and total assets in the previous five years; BIG4 is a dummy variable that equals 1 if the firm is audited by one of the Big 4 audit firms, zero otherwise. Both firm and year fixed effects are used. T-statistics robust to heteroscedasticity and clustered by firm are reported in the square brackets. *, ** and *** denote significance levels at 10%, 5% and 1% respectively.

Table 7: Loss of Seeking Alpha coverage

VARIABLES	Baseline SYNCH	Exogenous drop SYNCH	Baseline CORRE	Exogenous drop CORRE
L_SA	-0.187***	-0.185***	-2.330***	-2.315***
	[-11.963]	[-11.852]	[-14.323]	[-14.258]
L_SA_DROP		0.162**		1.875*
		[2.033]		[1.871]
LNUM	0.112***	0.112***	1.074***	1.073***
	[4.573]	[4.570]	[4.868]	[4.862]
SIZE	0.786***	0.786***	5.848***	5.847***
	[32.231]	[32.239]	[27.155]	[27.178]
LMB	0.044**	0.044**	0.311*	0.308*
	[2.236]	[2.222]	[1.694]	[1.680]
LEV	-0.219*	-0.218*	-0.726	-0.709
	[-1.831]	[-1.821]	[-0.724]	[-0.707]
ROA	-0.001	-0.001	4.646***	4.651***
	[-0.081]	[-0.082]	[6.692]	[6.699]
NIND	-0.166	-0.166	-1.021	-1.029
	[-1.544]	[-1.548]	[-1.121]	[-1.130]
HERFSALE	0.867	0.866	-6.266	-6.260
	[1.488]	[1.487]	[-1.145]	[-1.144]
STDROA	-0.025	-0.025	0.242	0.244
	[-1.188]	[-1.187]	[0.242]	[0.245]
BIG4	-0.077	-0.077	0.390	0.394
	[-1.234]	[-1.232]	[0.792]	[0.799]
Observations	108,333	108,333	112,346	112,346
firm fixed effects	Yes	Yes	Yes	Yes
time fixed effects	Yes	Yes	Yes	Yes
Adj. R-squared	0.665	0.665	0.675	0.675

Table 7 presents the effect of Seeking Alpha coverage on stock return co-movement using quarterly data. SYNCH, stock price synchronicity, is defined as the log-transformation of the adjusted R^2 of the firm-quarter estimation regressing daily stock return on daily market and industry level return; CORRE is the percentage points of time-series Pearson correlation coefficient between daily firm return and daily market returns; L_SA is the natural log of 1 plus the number of single-ticker Seeking Alpha articles for a firm in the quarter; L_SA_DROP is the natural log of 1 plus the number of exogenous drop of single-ticker Seeking Alpha articles for a firm in the quarter caused by influential authors; LNUM is the natural log of 1 plus the number of analysts following the firm for the previous fiscal year; SIZE is the natural log of the firm's market capitalisation at the end of the last fiscal year; LMB is the natural log of market capitalisation scaled by the book value of equity at the end of the last fiscal year; LEVERAGE is total long-term debt scaled by total assets at the end of the last fiscal year; ROA is income before extraordinary items divided by total assets at the end of the last fiscal year; NIND is the natural log of the number of firms in the industry to which firm i belongs at the end of the last fiscal year; HERFSALE is the sum of the squared terms of the proportion of a firm's revenue to total revenue in the industry at the end of the last fiscal year; STDROA is the standard deviation of the ratio between income before extraordinary items and total assets in the previous five years; BIG4 is a dummy variable that equals 1 if the firm is audited by one of the Big 4 audit firms, zero otherwise. T-statistics robust to heteroscedasticity and clustered by firm are reported in the square brackets. *, ** and *** denote significance levels at 10%, 5% and 1% respectively.

Table 8: “Day of the Week” effect

VARIABLES	SYNCH	CORRE
L_SA_WEEKENDS	-0.079***	-1.204***
	[-3.529]	[-5.397]
L_SA_WEEKDAY	-0.167***	-2.048***
	[-9.963]	[-12.239]
LNUM	0.112***	1.031***
	[4.567]	[4.645]
SIZE	0.786***	6.193***
	[32.228]	[30.549]
LMB	0.044**	0.139
	[2.221]	[0.838]
LEV	-0.219*	-0.975
	[-1.829]	[-0.961]
ROA	-0.001	0.103*
	[-0.085]	[1.895]
NIND	-0.167	-1.184
	[-1.558]	[-1.309]
HERFSALE	0.865	-6.036
	[1.485]	[-1.324]
STDROA	-0.025	-0.211
	[-1.183]	[-1.534]
BIG4	-0.077	0.329
	[-1.233]	[0.658]
Observations	108,333	112,346
firm fixed effects	Yes	Yes
time fixed effects	Yes	Yes
Adj. R-squared	0.665	0.674

Table 8 presents the effect of Seeking Alpha coverage on stock return co-movement using quarterly data. SYNCH, stock price synchronicity, is defined as the log-transformation of the adjusted R^2 of the firm-quarter estimation regressing daily stock return on daily market and industry level return; CORRE is the percentage points of time-series Pearson correlation coefficient between daily firm return and daily market returns; L_SA_WEEKENDS is the natural log of 1 plus the number of single-ticker Seeking Alpha articles posted on Friday, Saturday and Sunday for a firm in the quarter; L_SA_WEEKDAY is the natural log of 1 plus the number of single-ticker Seeking Alpha articles posted from Monday to Thursday for a firm in the quarter; LNUM is the natural log of 1 plus the number of analysts following the firm for the previous fiscal year; SIZE is the natural log of the firm’s market capitalisation at the end of the last fiscal year; LMB is the natural log of market capitalisation scaled by the book value of equity at the end of the last fiscal year; LEVERAGE is total long-term debt scaled by total assets at the end of the last fiscal year; ROA is income before extraordinary items divided by total assets at the end of the last fiscal year; NIND is the natural log of the number of firms in the industry to which firm i belongs at the end of the last fiscal year; HERFSALE is the sum of the squared terms of the proportion of a firm’s revenue to total revenue in the industry at the end of the last fiscal year; STDROA is the standard deviation of the ratio between income before extraordinary items and total assets in the previous five years; BIG4 is a dummy variable that equals 1 if the firm is audited by one of the Big 4 audit firms, zero otherwise. T-statistics robust to heteroscedasticity and clustered by firm are reported in the square brackets. *, ** and *** denote significance levels at 10%, 5% and 1% respectively.

Table 9: Incorporation of future earnings into current stock price

VARIABLES	4(a) Ret1	4(b) Ret2	Wald Test of coefficient difference
EARNING	2.186***	5.214***	
	[10.474]	[16.949]	
L_SA	0.051***	0.087***	
	[7.995]	[8.565]	
EARNING*L_SA	0.012***	0.056***	0.044***
	[3.342]	[6.972]	[23.07]
LNUM	0.056***	0.075***	
	[12.006]	[10.523]	
EARNING*LNUM	-0.043***	-0.104***	
	[-5.925]	[-7.421]	
SIZE	-0.077***	-0.152***	
	[-25.274]	[-34.164]	
EARNING*SIZE	0.041***	0.077***	
	[5.558]	[6.393]	
LMB	-0.160***	-0.222***	
	[-29.072]	[-27.075]	
EARNING*LMB	0.015***	0.022***	
	[3.053]	[3.061]	
LEVERAGE	0.337***	0.572***	
	[12.491]	[14.083]	
EARNING*LEVERAGE	0.961***	1.926***	
	[7.209]	[9.711]	
ROA	-0.042***	-0.115***	
	[-3.900]	[-5.938]	
EARNING*ROA	0.000	0.001***	
	[0.179]	[3.027]	
NIND	-0.001	0.009	
	[-0.110]	[1.076]	
EARNING* NIND	-0.373***	-0.854***	
	[-11.926]	[-18.372]	
HERFSALE	-0.081	0.060	
	[-1.002]	[0.483]	
EARNING*HERFSALE	0.356	-3.135***	
	[0.594]	[-3.482]	
STDROA	0.054***	0.057***	
	[5.019]	[4.460]	
EARNING*STDROA	-0.001***	-0.005***	
	[-3.221]	[-5.861]	
BIG4	0.133***	0.321***	
	[11.979]	[19.500]	
EARNING*BIG4	0.042*	0.112***	
	[1.764]	[3.237]	
CONSTANT	1.506***	1.955***	
	[39.930]	[34.206]	

Observations	34,866	34,866
Adj. R-squared	0.024	0.086

Table 9 presents the results on whether firms with higher Seeking Alpha coverage incorporate future earnings into their prices more efficiently. We estimate the system of Equations 10(a) and 10(b) using Seemingly Unrelated Regression (SUR). RET1 is the return from the previous year and RET2 is the previous two years' return; EARNINGS is the income before extraordinary items divided by total assets; L_SA is the natural log of 1 plus the number of single-ticker Seeking Alpha articles for a firm in the year; LNUM is the natural log of 1 plus the number of analysts following the firm for the previous fiscal year; SIZE is the natural log of the firm's market capitalisation at the end of the last fiscal year; LMB is the natural log of market capitalisation scaled by the book value of equity at the end of the last fiscal year; LEVERAGE is the total long-term debt scaled by total assets at the end of the last fiscal year; ROA is income before extraordinary items divided by total assets at the end of the last fiscal year; NIND is the natural log of the number of firms in the industry to which firm i belongs at the end of the last fiscal year; HERFSALE is the sum of the squared terms of the proportion of a firm's revenue to total revenue in the industry at the end of the last fiscal year; STDROA is the standard deviation of the ratio between income before extraordinary items and total assets in the previous five years; BIG4 is a dummy variable that equals 1 if the firm is audited by one of the Big 4 audit firms, zero otherwise. We report the Wald test results on the difference of coefficients on EARNINGS*L_SA between two equations. *, ** and *** denote significance levels at 10%, 5% and 1% respectively.

Table 10: Analysis comparing post-2011 and pre-2011 period

VARIABLES	SYNCH	CORRE
L_SA	0.198	-0.557
	[1.503]	[-0.407]
POST	-0.207	1.247
	[-0.579]	[0.241]
L_SA*POST	-0.140***	-3.652***
	[-3.720]	[-3.056]
LNUM	0.187***	2.377***
	[11.175]	[8.801]
SIZE	0.584***	6.338***
	[34.502]	[26.656]
LMB	-0.383***	-3.908***
	[-8.430]	[-8.464]
LEV	0.470***	3.913**
	[4.685]	[2.477]
ROA	-0.174	0.222
	[-1.405]	[0.130]
NIND	-0.021	0.126
	[-0.737]	[0.317]
HERFSALE	1.227***	7.279*
	[3.155]	[1.689]
STDROA	0.583***	6.527***
	[3.601]	[2.754]
BIG4	0.233***	4.522***
	[4.263]	[5.784]
CONSTANT	-6.374***	-2.850
	[-20.076]	[-0.776]
Observations	39,568	39,568
Firm clustering	Yes	Yes
Time clustering	Yes	Yes
Adj. R-squared	0.449	0.393

Table 10 presents results comparing the effect of Seeking Alpha coverage on co-movement between post-2011 and pre-2011 period. The sample contains 39,568 firm-year observations over the period 2004–2014. SYNCH, stock price synchronicity, is defined as the log-transformation of the adjusted R^2 of the firm-year estimation regressing weekly stock return on weekly market- and industry-level returns; CORRE is the percentage points of time-series Pearson correlation coefficient between weekly firm returns and the weekly market returns; L_SA is the natural log of 1 plus the number of single-ticker Seeking Alpha articles for a firm in the year; POST is a time dummy which equals to 1 for observations after 2011 and 0 otherwise; LNUM is the natural log of 1 plus the number of analysts following the firm for the previous fiscal year; SIZE is the natural log of the firm's market capitalisation at the end of the last fiscal year; LMB is the natural log of market capitalisation scaled by the book value of equity at the end of the last fiscal year; LEVERAGE is the total long-term debt scaled by total assets at the end of the last fiscal year; ROA is income before extraordinary items divided by total assets at the end of the last fiscal year; NIND is the natural log of the number of firms in the industry to which firm i belongs at the end of the last fiscal year; HERFSALE is the sum of the squared terms of the proportion of a firm's revenue to total revenue in the industry at the end of the last fiscal year; STDROA is the standard deviation of the ratio between income before extraordinary items and total assets in the previous five years; BIG4 is a dummy variable that equals one if the firm is audited by one of the Big 4 audit firms, zero otherwise. T-statistics robust to heteroscedasticity and clustered by firm and year are reported in the square brackets. *, ** and *** denote significance levels at 10%, 5% and 1% respectively.